Nonlinear Unknown Input Observability: The General Analytic Solution

Agostino Martinelli[1]

April 12, 2017

[1]A. Martinelli is with INRIA Rhone Alpes, Montbonnot, France e-mail: agostino.martinelli@inria.fr

If you have read this book and wish to be included in a mailing list (or other means of communication) that I maintain on the subject, then send e-mail To: agostino.martinelli@inria.fr

Contents

Abstract

This book provides the analytic solution of a fundamental open problem in control theory. It provides the general analytic criterion to obtain the state observability of nonlinear systems in presence of multiple unknown inputs. This problem, called the Unknown Input Observability (UIO) problem, was introduced in the seventies. This book provides and illustrates its general analytic solution. As for the observability rank condition (that is the analytic criterion to obtain the state observability of nonlinear systems in the case without unknown inputs), the analytic criterion in presence of unknown inputs is based on the computation of the observable codistribution by a recursive and convergent algorithm. The first part of the book provides this algorithm together with its convergence criterion. The algorithm is unexpectedly simple and can be easily and automatically applied to nonlinear systems driven by both known and unknown inputs, independently of their complexity and type of nonlinearity. Very surprisingly, the complexity of the overall analytic criterion is comparable to the complexity of the standard method to check the state observability in the case without unknown inputs (i.e., the observability rank condition). Given any nonlinear system characterized by any type of nonlinearity, driven by both known and unknown inputs, the state observability is obtained automatically, i.e., without human intervention (e.g., by the usage of a very simple code that uses symbolic computation). This is a fundamental practical (and unexpected) advantage. On the other hand, the analytic derivations and all the proofs necessary to analytically derive the algorithm and its convergence properties and to prove their general validity are very complex and they are extensively based on an ingenious analogy with the theory of General Relativity. In practice, these derivations largely use Ricci calculus with tensors (in particular, we largely adopt the Einstein notation to achieve notational brevity). The analytic criterion here derived is the follow up of the analytic criterion given in [43], that is the analytic solution of the UIO problem for nonlinear driftless systems in presence of a single unknown input. The new general analytic criterion accounts for the presence of a drift and the presence of multiple unknown inputs. Very interestingly, in the aforementioned analogy with the theory of General Relativity, the presence of the drift term corresponds to the presence of a time dimension in relativity and the presence of unknown inputs corresponds to the space dimension in relativity (i.e., in the aforementioned analogy, the space dimension of relativity equals the number of unknown inputs). In this sense, the solution in [43], which holds for driftless systems and with a single unknown input, corresponds to the trivial case of a space-time frozen with respect to time and with a single spatial dimension (this is the reason why the derivation of the solution in [43] did not require the use of Ricci calculus). The analytic criterion is illustrated by checking the observability of several nonlinear systems driven by multiple known inputs and multiple unknown inputs, ranging from planar robotics up to advanced nonlinear systems. In particular, the last applications are in the framework of visual-inertial sensor fusion. For this problem, the application of the analytic criterion provides a remarkable and amazing result, which could have relevance on the problem of visual-vestibular integration for self-motion perception in neuroscience.

Keywords: Nonlinear observability; Unknown Input Observability; Observability Rank Condition; Observable Codistribution; Unknown Input Observer; Robotics Applications; General Relativity; Visual-Inertial Sensor Fusion; Self-motion perception; Neuroscience

Chapter 1

Introduction

State observability is a necessary condition that a state must satisfy to be estimated. Checking the system observability is a fundamental step in state estimation. It has become praxis in many application fields to provide an observability analysis prior to solving an estimation problem (e.g., in robotics [6, 22, 24, 25, 33, 47, 48, 49, 51], in visual-inertial sensor fusion [19, 20, 21, 28, 29, 32, 34, 38, 39, 44, 46], in sensor calibration [7, 15, 35, 36, 45]). Investigating the observability properties is very simple in the linear case. Unfortunately, real systems are very rarely characterized by linearity.

The control theory community has developed the analytic tool necessary to check the state observability for nonlinear systems provided that all the system inputs are known. This is the observability rank condition introduced by Herman and Krener in 1977 [18]. In accordance with this criterion, it is possible to derive all the observability properties of a nonlinear system by performing automatic computation, independently of the complexity and the type of nonlinearity. On the other hand, in many real scenarios, one or more disturbances can dramatically impact the system dynamics. A disturbance can be considered as an unknown input. Its presence can dramatically affect the observability properties of the state. This is for instance the case of a drone that operates in presence of wind. The wind is in general unknown, time-variant and acts on the system dynamics as an unknown input.

The problem of deriving an analytic tool able to determine the observability properties in presence of unknown inputs is known as the *Unknown Input Observability* (UIO) problem. This problem was introduced and firstly investigated by the automatic control community in the seventies [3, 5, 14, 50]. In particular, Basile and Marro provided the solution of this problem in the linear case [3]. In many application fields, most of the systems are characterized by nonlinear dynamics, even in very simple cases (e.g. in planar robotics). Additionally, the presence of disturbances cannot be ignored in many cases and can significantly affect the observability properties.

The UIO problem in the nonlinear case has recently been investigated and partial solutions have recently been proposed [4, 39, 40, 41]. These solutions only provide sufficient conditions for the state observability. They are based on a suitable state extension.

A great effort has been devoted to design observers for both linear and nonlinear systems in presence of UI, in many cases in the context of fault diagnosis, e.g., [1, 2, 8, 9, 10, 11, 12, 13, 16, 17, 23, 30, 31, 52]. In some of these works, interesting conditions for the existence of an unknown input observer were also derived. On the other hand, these conditions have the following limitations:

- They refer to a restricted class of systems since they are often characterized by linearity (or

some specific type of nonlinearity) with respect to the state in some of the functions that characterize the dynamics[1] and/or the system outputs. No condition refers to any type of nonlinearity with respect to the state in the aforementioned functions.

- They cannot be implemented automatically, i.e., without human intervention (e.g., by the usage of a simple code that adopts a symbolic computation tool).

These limitations do not affect the observability rank condition in [18, 27]. However, this condition cannot be used in presence of Unknown Input (UI). The extension of the observability rank condition to the UI case is a simple analytic condition able to provide the state observability in presence of UI that does not encounter the two limitations of above. This is the solution of the UIO problem. The goal of this book is to provide this extension. This can be considered the natural extension of the observability rank condition introduced in [18] to the case when the dynamics are also driven by unknown inputs.

Very recently, the analytic solution of the UIO problem in the case of a single unknown input has been introduced [42, 43]. This solution holds for any dynamics nonlinear in the state and linear in the inputs (both known and unknown). This book provides the complete analytic solution of the UIO problem in the nonlinear case. In chapter 2 we define the class of systems for which we provide the solution. This class is very general and basically includes any nonlinear system affine in the inputs (both known and unknown) and any number of known and unknown inputs.

In [18, 27] the observability properties of a nonlinear system are obtained by computing the observable codistribution. The computation of this codistribution is the core of the observability rank condition introduced in [18]. In order to deal with the case of unknown inputs, we need to derive a new algorithm able to generate the observable codistribution. In chapter 3 we remind the reader the algorithm to compute the observable codistribution in the case without disturbances, together with some basic properties that characterize its convergence (section 3.1). Then, in sections 3.2 and 3.3, we introduce the new algorithm that generates the entire observable codistribution in presence of disturbances, together with some basic properties that characterize its convergence. The solution of the UIO problem is summarized in chapter 4 and it is illustrated in chapter 5 by checking the observability of several nonlinear systems driven by multiple known inputs and multiple unknown inputs, ranging from planar robotics up to advanced nonlinear systems. In particular, the last applications are in the framework of visual-inertial sensor fusion. For this problem, the application of the analytic criterion provides a remarkable and intriguing result, which could impact the current research conducted by the neuroscience community about the visual-vestibular integration for self-motion perception (see sections 5.6 and 5.7 for a detailed discussion about this application). Note that the method is very powerful and can be used to automatically obtain the observability properties of any system that satisfies (2.1), i.e., independently of the state dimension (intuitive reasoning becomes often prohibitive for systems characterized by high-dimensional states), independently of the type of nonlinearity and in general independently of the system complexity. To this regard, note that the method can be implemented automatically, by using a simple symbolic computation tool (in our examples, in most of cases we executed the computation manually, and, in the hardest cases, we simply used the symbolic toolbox of MATLAB). In chapter 6 we provide all the analytical derivations necessary to prove the validity of the results presented in chapter 3.

Throughout this book we often consider a simplified case before considering the general case. Specifically, in the simplified case, the dynamics are linear (and not affine) in the inputs (both known and unknown) and they are characterized by a single unknown input. This simplified

[1]These functions are the functions that appear in equation (2.1), i.e., $g^0(x), g^1(x), \cdots, g^{m_w}(x), f^1(x), \cdots, f^{m_u}(x)$.

case is precisely the case of a nonlinear driftless system with a single unknown input, for which the analytic solution was previously published [42, 43].

Chapter 2

System definition

We will refer to a nonlinear control system with m_u known inputs ($u \triangleq [u_1, \cdots, u_{m_u}]^T$) and m_w unknown inputs or disturbances ($w \triangleq [w_1, \cdots, w_{m_w}]^T$). The state is the vector $x \in M$, with M an open set of $\mathbb{R}^n$. We assume that the dynamics are nonlinear with respect to the state and affine with respect to the inputs (both known and unknown). Finally, for the sake of simplicity, we will refer to the case of a single output y and we provide the extension to multiple outputs, which is trivial, separately. This will allow us to avoid the introduction of a further index. Our system is characterized by the following equations:

$$\begin{cases} \dot{x} = g^0(x) + \sum_{i=1}^{m_u} f^i(x) u_i + \sum_{j=1}^{m_w} g^j(x) w_j \\ \\ y = h(x) \end{cases} \tag{2.1}$$

where $g^0(x)$, $f^i(x)$, $i = 1, \cdots, m_u$, and $g^j(x)$, $j = 1, \cdots, m_w$, are vector fields in M and the function $h(x)$ is a scalar function defined on the open set M. Finally, we assume that the unknown inputs $w_1, \cdots, w_{m_w}$ are analytic functions of time.
Throughout this book, in the case when $m_u = 1$, we denote by $f(x)$ the vector field $f^1(x)$ and by u the known input u_1. Similarly, in the case $m_w = 1$, we denote by $g(x)$ the vector field $g^1(x)$ and by w the unknown input w_1.

As mentioned in the introduction, throughout this book, we often consider the simpler system characterized by $m_w = 1$ and dynamics linear (and not affine) in the inputs (i.e., without the term g^0). In other words, the system characterized by the following equations:

$$\begin{cases} \dot{x} = \sum_{i=1}^{m_u} f^i(x) u_i + g(x) w \\ y = h(x) \end{cases} \tag{2.2}$$

In particular, we often provide the results for the system characterized by (2.2) before considering the general case characterized by (2.1). We often refer to this system as to the case when $m_w = 1$ and $g^0 = 0$. Note that this simplified case is precisely the case previously published [42, 43].

Chapter 3

Observable Codistribution

As for the observability rank condition, the analytic method to investigate the observability properties of the system in (2.1), is obtained by computing the observable codistribution[1]. In this chapter we provide the algorithm to compute this codistribution. For educational purposes, we start by reminding the reader the standard algorithm that generates the observable codisribution in absence of unknown inputs (section 3.1). Then, in section 3.2, we provide the algorithm for the system in (2.2). Finally, in 3.3, we provide the algorithm in the general case. For each algorithm we also provide the convergence criterion. Note that this chapter directly provides the analytic results. All the analytic proofs will be provided later, in chapter 6.

[1]The reader non-familiar with the concept of *distribution*, as it is used in [27], should not be afraid by the term *distribution* and the term *codistribution*. Very simply speaking, a distribution is a vector space defined on M (our set in $\mathbb{R}^n$ where the system is defined). In particular, this vector space changes by moving on M. This vector space is in fact the span of a set of vector functions (vector fields) defined on M. A codistribution is the dual of a distribution. Very simply speaking (and this is enough to understand this manuscript) a distribution is generated by a set of column vectors. A codistribution is generated by a set of line vectors. All these vectors are vector functions (i.e., they depend on the point $x \in M$) and they have the same dimension of x.

3.1 Observable codistribution in the case without unknown inputs

We consider the system in (2.1) when $m_w = 0$ (all the inputs are known). We will denote with the symbol $\mathcal{D}$ the differential with respect to the state x. For instance, if $x = [x_1,\ x_2]^T$ and $h = x_1 + x_2^2$, we have[2]: $\mathcal{D}h = \mathcal{D}x_1 + 2x_2\mathcal{D}x_2$.

For a given codistribution Ω and a given vector field $f = f(x)$ (both defined on the open set M), we denote by $\mathcal{L}_f\Omega$ the codistribution whose covectors are the Lie derivatives along f of the covectors in Ω. We remind the reader that the Lie derivative of a scalar function $h(x)$ along the vector field $f(x)$ is defined as follows:

$$\mathcal{L}_f h \triangleq \frac{\partial h}{\partial x} f$$

which is the product of the row vector $\frac{\partial h}{\partial x}$ with the column vector f. Hence, it is a scalar function. Additionally, by definition of Lie derivative of covectors, we have: $\mathcal{L}_f\mathcal{D}h = \mathcal{D}\mathcal{L}_f h$.

Finally, given two vector spaces V_1 and V_2, we denote by $V_1 + V_2$ their sum, i.e., the span of all the generators of both V_1 and V_2.

The observable codistribution is generated by the following recursive algorithm (see [18] and [27]):

Algorithm 1 Observable codistribution in the case $m_w = 0$

1. $\Omega_0 = span\{\mathcal{D}h\}$;
2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f^i}\Omega_{m-1} + \mathcal{L}_{g^0}\Omega_{m-1}$

In presence of multiple outputs, we only need to add to the codistribution Ω_0, the span of the differentials of the remaining outputs. In [27] it is proven that this algorithm converges. In particular, it is proven that it has converged when $\Omega_m = \Omega_{m-1}$. From this, it is easy to realize that the convergence is achieved in at most $n - 1$ steps[3].

3.2 Observable codistribution in the case without drift and with a single unknown input

We now provide the new algorithm that generates the observable codistribution in presence of unknown inputs. For the clarity sake, we start by providing the algorithm for the system characterized by equation (2.2). We will denote by $L_g^1 = L_g^1(x)$ the first order Lie derivative of the function $h(x)$ along the vector field $g(x)$, i.e.,

$$L_g^1 \triangleq \mathcal{L}_g h \tag{3.1}$$

The analytic computation of the observable codistribution is based on the assumption that $L_g^1 \neq 0$ on a given neighbourhood of x_0. In appendix A, we introduce the concept of *canonic*

[2]The span of the differentials of a set of scalar functions is a codistribution. The reader non familiar with the theory of distributions can simply consider the differential as the gradient operator. The gradient of a scalar function is a line vector. For instance, if $x = [x_1,\ x_2]^T$ and $h = x_1 + x_2^2$, we obtain for its gradient the line vector function $[1,\ 2x_2]$. Later, in chapter 5, we adopt this representation. According to this, a codistribuion will be the span of a set of line vectors and a covector (i.e., an element of a codistribution) will be a line vector.

[3] This is a consequence of lemmas 1.9.1, 1.9.2 and 1.9.6 in [27].

form with respect to the unknown inputs. For the case $m_w = 1$ (dealt in section A.1), we show that the system is in canonic form with respect to the unknown input w, if either $L^1_g \neq 0$ or it is possible to redefine the output, without altering the system observability properties, such that the Lie derivative of the new output along g does not vanish[4]. Finally, if a system characterized by $m_w = 1$ is not in canonic form, the unknown input is spurious (i.e., it does not affect the observability properties). For these reasons, we can assume that $L^1_g \neq 0$.

Before introducing the new algorithm that generates the entire observable codistribution, we introduce a new set of vector fields ${}^i\phi_m \in \mathbb{R}^n$ ($i = 1, \cdots, m_u$ and for any integer m). They are obtained recursively by the following algorithm:

Algorithm 2

1. ${}^i\phi_0 = f^i$;
2. ${}^i\phi_m = \frac{[{}^i\phi_{m-1},\ g]}{L^1_g}$

where the parenthesis $[\cdot, \cdot]$ denote the Lie bracket of vector fields, defined as follows:

$$[a,\ b] \triangleq \frac{\partial b}{\partial x} a(x) - \frac{\partial a}{\partial x} b(x)$$

In other words, for each $i = 1, \cdots, m_u$, we have one new vector field at every step of the algorithm. Throughout this book, in the case when $m_u = 1$, we denote by ϕ_m the vector field ${}^1\phi_m$.

We are now ready to provide the algorithm that generates the entire observable codistribution. It is the following:

Algorithm 3 Observable codistribution in the case $m_w = 1$ and $g^0 = 0$

1. $\Omega_0 = span\{\mathcal{D}h\}$;
2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f^i}\Omega_{m-1} + \mathcal{L}_{\frac{g}{L^1_g}}\Omega_{m-1} + \sum_{i=1}^{m_u} span\left\{\mathcal{L}_{{}^i\phi_{m-1}}\mathcal{D}h\right\}$

In presence of multiple outputs, we only need to add to the codistribution Ω_0, the span of the differentials of the remaining outputs. Note that, in presence of multiple outputs, the function L^1_g is still a scalar function since it is still defined by using a single output. The result is independent of the chosen output (provided that L^1_g does not vanish[5]).

In section 6.2.2 we investigate the convergence properties of algorithm 3. We consider first the case of a single known input (i.e., $m_u = 1$) and then the results are easily extended to the case of multiple inputs ($m_u > 1$) in section 6.2.3. We prove that algorithm 3 converges and we also provide the analytic criterion to check that the convergence has been attained. This proof and the convergence criterion cannot be the same that hold for algorithm 1, because of the last term that appears in the recursive step[6], i.e., the term $\sum_{i=1}^{m_u} span\left\{\mathcal{L}_{{}^i\phi_{m-1}}\mathcal{D}h\right\}$ (the special case when, the contribution due to this last term is included in the other terms, is considered separately by lemma 6). In general, the criterion to establish that the convergence has been attained is not simply obtained by checking if $\Omega_{m+1} = \Omega_m$. Deriving the new analytic criterion

[4]The new output is selected from the space of functions $\mathcal{F}$, defined, in this case, as the space that contains h and its Lie derivative up to any order along the vector fields $f^1, \cdots, f^{m_u}$.

[5]The case when $L^1_g = 0$ for all the outputs is dealt in appendix A.

[6]The convergence criterion of algorithm 1 is a consequence of the fact that, all the terms that appear in the recursive step of algorithm 1, are the Lie derivative of the codistribution at the previous step, along fixed vector fields (i.e., vector fields that remain the same at each step of the algorithm). This is not the case for the last term in the recursive step of algorithm 3.

is not immediate. It requires to derive the analytic expression that describes the behaviour of the last term in the recursive step. This fundamental equation is provided in chapter 6 and it is the equation (6.17). The analytic derivation of this equation allows us to detect the key quantity that governs the convergence of algorithm 3, in particular regarding the contribution due to the last term in the recursive step. This key quantity is the following scalar:

$$\tau \triangleq \frac{\mathcal{L}_g^2 h}{(L_g^1)^2} \tag{3.2}$$

We prove (see lemma 10 in chapter 6) that, in general, it exists m' such that $\mathcal{D}\tau \in \Omega_{m'}$ (and therefore $\mathcal{D}\tau \in \Omega_m \ \forall m \geq m'$). Additionally, we prove that the convergence of the algorithm has been reached when $\Omega_{m+1} = \Omega_m$, $m \geq m'$ and $m \geq 2$ (theorem 2). We also prove that the required number of steps is at most $n+2$.

In section 6.2.1 it is also shown that the computed codistribution is the entire observable codistribution. Also in this case, the proof is given by first considering the case of a single known input (see theorem 1) and then, its validity is extended to the case of multiple inputs in section 6.2.3. Note that this proof is based on the assumption that the unknown input (w) is a differentiable function of time, up to a given order (the order depends on the specific case).

Algorithm 3 differs from the standard algorithm 1 because of the following reasons:

- In the recursive step, the vector field that corresponds to the unknown input (i.e., g) must be rescaled by dividing by L_g^1.
- The recursive step also contains the sum of the contributions $\sum_{i=1}^{m_u} \mathcal{L}_{^i\phi_{m-1}} \mathcal{D}h$. In other words, we need to compute the Lie derivatives of the differential of the output along the vector fields obtained through the recursive algorithm 2.
- The convergence of algorithm 3 is achieved in at most $n+2$ steps, instead of $n-1$ steps (in the special case dealt by lemma 6, this upper bound is $n-1$ for both cases).
- When $\Omega_m = \Omega_{m-1}$ algorithm 1 has converged. For algorithm 3, we also need to check that $\mathcal{D}\tau \in \Omega_m$ and $m \geq 2$ (with the exception of the special case dealt by lemma 6).

3.3 Observable codistribution in the general case

In the sequel, when dealing with the case $m_w > 1$, it is very useful, for notational brevity, to adopt the Einstein notation, where m_w plays the role of the spatial dimension in General relativity and the presence of the term g^0 corresponds to the time dimension. Specifically, Latin indexes will take the values $1, 2, \cdots, m_w$ while Greek indexes the values $0, 1, 2, \cdots, m_w$. According to this, the dynamics in (2.1) can be written as follows: $\dot{x} = g^0 + \sum_{i=1}^{m_u} f^i u_i + g^j w_j$ (note that, according to this notation, the sum on j is omitted). In general, according to this notation, an index that is summed over is a summation index, in this case j. It is also called a dummy index since any symbol can replace j without changing the meaning of the expression provided that it does not collide with index symbols in the same term. When the dummy index is Latin, the sum is from 1 to m_w. When it is Greek, the sum is from 0 to m_w. In addition, let us suppose that we have a tensor equation like $\Gamma_k^\alpha = 0$. To specify that this equation holds for $\alpha = 0, 1, \cdots, m_w$ and $k = 1, \cdots, m_w$, we simply write $\Gamma_k^\alpha = 0$, $\forall \alpha, k$. Similarly, if we want to refer to the components of the tensor Γ_k^α, for $\alpha = 0, 1, \cdots, m_w$ and $k = 1, \cdots, m_w$, we simply write Γ_k^α, $\forall \alpha, k$.

We start by selecting a set of m_w scalar functions that we denote by $h_1, \ h_2, \cdots, h_{m_w}$. These functions are selected from the output and any order Lie derivative of the output only along

the vector fields that correspond to the known inputs, i.e., $f^i(x)$, $i = 1, \cdots, m_u$. In the case of multiple outputs, we can use different outputs for this selection[7]. We define the two-index tensor:

$$\mu^i_j \triangleq \mathcal{L}_{g^i} h_j, \qquad \forall i, j \tag{3.3}$$

Note that this tensor is of type $(1, 1)$, i.e., it has one upper index and one lower index. The analytic computation of the observable codistribution is based on the assumption that the tensor μ is non singular on a given neighbourhood of x_0. In appendix A, we introduce the concept of *canonic form* with respect to the unknown inputs. The system is in its canonic form with respect to its unknown inputs if it is possible to select m_w scalar functions (as previously specified) such that the corresponding tensor μ is non singular. In appendix A, we show that, for a system that is not in its canonic form, it is possible either to find a finite number of local coordinates changes in the space of the unknown inputs and their time derivatives up to a given order, such that the canonic form is achieved, or to show that some of the unknown inputs are spurious (i.e., they do not affect the observability properties). In this latter case, it is possible to write the dynamics in (2.1) with a number of unknown inputs smaller than m_w and in canonic form with respect to these new unknown inputs. Hence, we can assume that the tensor μ is non singular.

We denote by ν the inverse of μ. In other words, according to the Einstein notation, we have:

$$\mu^i_k \nu^k_j = \delta^i_j, \qquad \forall i, j \tag{3.4}$$

where, in accordance with the Einstein notation, the dummy Latin index k is summed over $k = 1, \cdots, m_w$, and δ is the Kronecker tensor. In the case $m_w = 1$ (discussed in section 3.2) we have $\mu^1_1 = L^1_g$ and $\nu^1_1 = \frac{1}{L^1_g}$.
Note that the tensors μ and ν previously defined have both indexes Latin. We need to complete their definition by including the components with 0 index. First of all, we extend equation (3.4) to the following equation:

$$\mu^\alpha_\gamma \nu^\gamma_\beta = \delta^\alpha_\beta, \qquad \forall \alpha, \beta \tag{3.5}$$

where, in accordance with the Einstein notation, the dummy Greek index γ is summed over $\gamma = 0, 1, \cdots, m_w$. Hence, It suffices to complete the definition of only one of the two tensors. We define μ^0_i the Lie derivative of h_i along g^0, i.e.,

$$\mu^0_i \triangleq \mathcal{L}_{g^0} h_i, \qquad \forall i \tag{3.6}$$

The remaining components of μ will be actually provided for a specific coordinate setting (in the analogy with the theory of General Relativity, this setting corresponds to the choice of an inertial frame, which is locally always possible). We set:

$$\mu^0_0 = 1, \qquad \mu^i_0 = 0, \quad \forall i \tag{3.7}$$

In this coordinate setting, we trivially obtain from (3.5):

$$\nu^0_0 = 1, \qquad \nu^0_i = -\nu^j_i \mu^0_j, \;\; \nu^i_0 = 0, \quad \forall i \tag{3.8}$$

We introduce the following $m_w + 1$ vector fields $\hat{g}^\alpha$, $\forall \alpha$:

$$\hat{g}^\alpha \triangleq \nu^\alpha_\beta g^\beta \tag{3.9}$$

[7]In appendix A, the space of functions that consists of the outputs and their Lie derivatives, up to any order, along $f^1, \cdots, f^{m_u}$, is denoted by $\mathcal{F}$.

where, in accordance with the Einstein notation, the dummy Greek index β is summed over $\beta = 0, 1, \cdots, m_w$. Note that, when $m_w = 1$ and $g^0 = 0$, the previous equation defines the single vector field $\hat{g}^1 \triangleq \frac{g^1}{L_g^1}$, which is the vector field that appears in algorithm 3 (where we denoted $g \triangleq g^1$).

Finally, we introduce the following abstract operation that generalizes the Lie bracket.

Definition 1 (Lie Bracket along a set of vector fields) *Given a vector field ϕ, we call the Lie bracket of ϕ along the set of vector fields g^α, $\forall\alpha$, through the $(1,1)$ tensor ν, the following set of $m_w + 1$ vector fields:*

$$[\phi]^\alpha \triangleq \nu^\alpha_\beta [\phi, \ g^\beta], \qquad \forall\alpha \tag{3.10}$$

where, in accordance with the Einstein notation, the dummy Greek index β is summed over $\beta = 0, 1, \cdots, m_w$. Note that the previous operation generalizes the operation of Lie bracket. Specifically, in absence of the term g^0 and when $m_w = 1$ and $\nu^1_1 = 1$, it reduces to the Lie bracket of ϕ along g^1. In the same setting, but when $\nu^1_1 = \frac{1}{L^1_g}$, we obtain the same operation that appears in the second line of algorithm 2.

Thanks to this new operation, we can express in compact form the general algorithm that provides the entire observable codistribution. We start by providing the analogous of algorithm 2. In this case, the new algorithm generates, at the m^{th} step, $(m_w + 1)^m$ vector fields for every $i = 1, \cdots, m_u$. We have:

Algorithm 4

1. ${}^i\phi_0 = f^i$;

2. ${}^i\phi_m^{\alpha_1, \cdots, \alpha_m} = [{}^i\phi_{m-1}^{\alpha_1, \cdots, \alpha_{m-1}}]^{\alpha_m}, \quad \forall\alpha_1, \cdots, \alpha_m$

which coincides with algorithm 2 when $m_w = 1$ and $g^0 = 0$. The algorithm that generates the entire observable codistribution is:

Algorithm 5 Observable codistribution in the case with multiple unknown inputs and $g^0 \neq 0$

1. $\Omega_0 = span\{\mathcal{D}h_l\}$

2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f^i}\Omega_{m-1} + \mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1} + \sum_{i=1}^{m_u} span\left\{\mathcal{L}_{{}^i\phi_{m-1}^{\alpha_1, \cdots, \alpha_{m-1}}}\mathcal{D}h_l\right\}$

Note that in the previous algorithm there are several dummy indexes $(l, \alpha, \alpha_1, \cdots, \alpha_{m-1})$ meaning that the corresponding codistibution must be summed for all the values that the indexes can take. For instance, the first equation of the algorithm is $\Omega_0 = \sum_{l=1}^{m_w} span\{\mathcal{D}h_l\}$ and the third contribution in the recursive step is $\sum_{\alpha=0}^{m_w} \mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1}$.

In presence of multiple outputs, we only need to add to the codistribution Ω_0, the span of the differentials of those outputs that have not been used in the selection of the functions $h_1, \cdots, h_{m_w}$.

In section 6.3.2 we investigate the convergence properties of algorithm 5. We consider first the case of a single known input (i.e., $m_u = 1$) and then the results are easily extended to the case of multiple known inputs ($m_u > 1$) in section 6.3.3. We prove that algorithm 5 converges and we also provide the analytic criterion to check that the convergence has been attained. As for algorithm 3, this proof and the convergence criterion cannot be the same that hold for algorithm 1, because

of the last term that appears in the recursive step[8], i.e., the term $\sum_{i=1}^{m_u} span \left\{ \mathcal{L}_{{}^i\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}} \mathcal{D}h_l \right\}$ (the special case when, the contribution due to this last term is included in the other terms, is considered separately by lemma 14). In general, the criterion to establish that the convergence has been attained is not simply obtained by checking if $\Omega_{m+1} = \Omega_m$. Deriving the new analytic criterion is a very laborious and demanding task. It requires to derive the analytic expression that describes the behaviour of the last term in the recursive step. This fundamental equation is provided in chapter 6 and it is the equation (6.37). The analytic derivation of this equation allows us to detect the key quantity that governs the convergence of algorithm 5, in particular regarding the contribution due to the last term in the recursive step. Its derivation was very troublesome since we did not a priori know that, instead of a scalar as for algortihm 3, the key quantity becomes a three-index tensor. Specifically, it is the following three-index tensor of type $(2,1)$:

$$\mathcal{T}_\gamma^{\alpha,\beta} \triangleq \nu_\eta^\beta (\mathcal{L}_{\hat{g}^\alpha} \mu_\gamma^\eta), \quad \forall \alpha, \beta, \gamma \tag{3.11}$$

where, in accordance with the Einstein notation, the dummy Greek index η is summed over $\eta = 0, 1, \cdots, m_w$. Note that, when $m_w = 1$ and $g^0 = 0$, this tensor has the single component $\mathcal{T}_1^{1,1}$, which coincides with the quantity τ defined in (3.2). In general, this tensor has $(m_w + 1) \times (m_w + 1) \times (m_w + 1)$ components. On the other hand, in our coordinate setting (for which equations (3.7) and (3.8) hold) it is immediate to obtain that $\mathcal{T}_0^{\alpha,\beta} = 0 \ \forall \alpha, \beta$. In other words, in this setting we can consider the lower index as a Latin index and the components of this tensor are $(m_w + 1) \times (m_w + 1) \times m_w$. Note that, in the rest of this book we will work in this coordinate setting. We prove (see lemma 18 in chapter 6) that, in general, it exists m' such that the differentials of all these components belong to $\Omega_{m'}$ (and therefore belong to $\Omega_m \ \forall m \geq m'$). Additionally, we prove that the convergence of the algorithm has been reached when $\Omega_{m+1} = \Omega_m$, $m \geq m'$ and $m \geq 2$. We also prove that the required number of steps is at most $n + 2$.

In section 6.3.1 it is also shown that the computed codistribution is the entire observable codistribution. Also in this case, the proof is given by first considering the case of a single known input and then, its validity is extended to the case of multiple inputs in section 6.3.3. Note that this proof is based on the assumption that the unknown inputs are differentiable functions of time, up to a given order (the order depends on the specific case).

[8] Again, we remind the reader that the convergence criterion of algorithm 1 is a consequence of the fact that, all the terms that appear in the recursive step of algorithm 1, are the Lie derivative of the codistribution at the previous step, along fixed vector fields (i.e., vector fields that remain the same at each step of the algorithm). This is not the case for the last term in the recursive step of algorithm 5.

Chapter 4

The Analytic Method

In this chapter, we outline all the steps to investigate the weak local observability at a given point x_0 of a nonlinear system characterized by (2.1). Basically, these steps are the steps necessary to compute the observable codistribution (i.e., the steps of algorithms 3 and 5 for the systems described by (2.2) and (2.1), respectively) and to prove that the differential of a given state component belongs to this codistribution. For the clarity sake, we distinguish the case $m_w = 1,\ g^0 = 0$ from the general case.

4.1 Method in the case without drift and with a single unknown input

Note that, in the trivial case analyzed by lemma 6, the method provided below simplifies, since we do not need to compute the quantity $\tau \left(= \frac{\mathcal{L}_g^2 h}{(\mathcal{L}_g^1 h)^2}\right)$, and we do not need to check that its differential belongs to the codistribution computed at every step of algorithm 3. In practice, we skip the steps 4 and 5 in the procedure below.

1. For the chosen x_0, compute $L_g^1 (= \mathcal{L}_g^1 h)$. In the case when $L_g^1 = 0$, choose another function in the space of functions $\mathcal{F}$ (defined as the space that contains h and its Lie derivative up to any order along the vector fields $f^1, \cdots, f^{m_u}$) such that its Lie derivative along g does not vanish[1].

2. Compute the codistribution Ω_0 and Ω_1 (at x_0) by using algorithm 3.

3. Compute the vector fields ${}^i\phi_m$ $(i = 1, \cdots, m_u)$ by using algorithm 2, starting from $m = 0$, to check if the considered system is in the special case dealt by lemma 6. In this trivial case, set $m' = 0$, use the recursive step of algorithm 3 to build the codistribution Ω_m for $m \geq 2$, and skip to step 6.

4. Compute $\tau \left(= \frac{\mathcal{L}_g^2 h}{(L_g^1)^2}\right)$ and $\mathcal{D}\tau$.

5. Use the recursive step of algorithm 3 to build the codistribution Ω_m for $m \geq 2$, and, for each m, check if $\mathcal{D}\tau \in \Omega_m$. Denote by m' the smallest m such that $\mathcal{D}\tau \in \Omega_m$.

[1]If the Lie derivative of any function in $\mathcal{F}$ vanishes, it means that the unknown input can be ignored to obtain the observability properties (the system is not canonic with respect to the unknown input, as shown in appendix A).

6. For each $m \geq m'$, check if $\Omega_{m+1} = \Omega_m$ and denote by $\Omega^* = \Omega_{m^*}$, where m^* is the smallest integer such that $m^* \geq m'$ and $\Omega_{m^*+1} = \Omega_{m^*}$ (note that $m^* \leq n+2$).
7. If the differential of a given state component (x_j, $j = 1, \cdots, n$) belongs to Ω^* (namely if $\mathcal{D}x_j \in \Omega^*$) on a given neighbourhood of x_0, then x_j is weakly locally observable at x_0. If this holds for all the state components, the state x is weakly locally observable at x_0. Finally, if the dimension of Ω^* is smaller than n on a given neighbourhood of x_0, then the state is not weakly locally observable at x_0.

4.2 Method in the general case

Note that, in the trivial case analyzed by lemma 14, the method provided below simplifies, since we do not need to compute the tensor $\mathcal{T}$ defined in (3.11), and we do not need to check that the differentials of its components belong to the codistribution computed at every step of algorithm 5. In practice, we skip the steps 5 and 6 in the procedure below.

1. Select a set of m_w scalar functions (h_1, $h_2, \cdots, h_{m_w}$) among the output (or the outputs) and any order Lie derivatives of the output (or the outputs) only along $f^i(x)$, $i = 1, \cdots, m_u$, such that the two index tensor $\mu^i_j \triangleq \mathcal{L}_{g^i} h_j$ defined in (3.3) is non singular. In the case is singular for any of the previous selections, apply the procedure provided in appendix A in order to set the system in canonic form. In the case the system cannot be set in canonic form, it means that some of the unknown inputs can be ignored to obtain the observability properties (note that the procedure detects these unknown inputs, automatically).
2. Complete the computation of μ by computing its components with 0 index (equations (3.6) and (3.7)). Compute the two index tensor ν (i.e., the inverse of μ) and the $m_w + 1$ vector fields: g^α defined in (3.9) ($\alpha = 0, 1, \cdots, m_w$).
3. Compute the codistribution Ω_0 and Ω_1 (at x_0) by using algorithm 5.
4. Compute the vector fields ${}^i\phi_m^{\alpha_1, \cdots, \alpha_m}$ ($i = 1, \cdots, m_u$, $\forall \alpha_1, \cdots, \alpha_m$) by using algorithm 4, starting from $m = 0$, to check if the considered system is in the special case dealt by lemma 14. In this trivial case, set $m' = 0$, use the recursive step of algorithm 5 to build the codistribution Ω_m for $m \geq 2$, and skip to step 7.
5. Compute the three-index tensor $\mathcal{T}$ defined in (3.11) and the differentials of all its components.
6. Use the recursive step of algorithm 5 to build the codistribution Ω_m for $m \geq 2$, and, for each m, check if $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m$, $\forall \alpha, \beta, k$. Denote by m' the smallest m such that $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m$, $\forall \alpha, \beta, k$.
7. For each $m \geq m'$ check if $\Omega_{m+1} = \Omega_m$ and denote by $\Omega^* = \Omega_{m^*}$, where m^* is the smallest integer such that $m^* \geq m'$ and $\Omega_{m^*+1} = \Omega_{m^*}$ (note that $m^* \leq n+2$).
8. If the differential of a given state component (x_j, $j = 1, \cdots, n$) belongs to Ω^* (namely if $\mathcal{D}x_j \in \Omega^*$) on a given neighbourhood of x_0, then x_j is weakly locally observable at x_0. If this holds for all the state components, the state x is weakly locally observable at x_0. Finally, if the dimension of Ω^* is smaller than n on a given neighbourhood of x_0, then the state is not weakly locally observable at x_0.

Chapter 5

Applications

We apply the method described in chapter 4 in order to investigate the observability properties of several nonlinear systems characterized by the equations given in (2.1). For educational purposes, we start by very simple examples where we analyze the observability properties of systems whose dynamics are affected by a single unknown input and are linear in the inputs. These examples will be discussed in sections 5.1 and 5.2. The dynamics of both these examples are the ones of the unicycle. In accordance with the unicycle dynamics, the motion is powered by two independent controls, which are the linear and the angular speed, respectively. In section 5.1 we consider the case when one of these two inputs is unknown and acts as an unknown input. In section 5.2 we consider the case when both these two inputs are known. However, the dynamics are also affected by an external unknown input.

Both these examples are deliberately very trivial in order to allow us to compare the analytic results provided by the proposed method with what we can expect by following intuitive reasoning.

In sections 5.4-5.7, we discuss a very important sensor fusion problem, that is the problem of fusing visual and inertial measurements. We start this discussion in the planar case. This allows us to investigate an example where the dynamics are affected by two unknown inputs and contain a drift term (g^0). In other words, the dynamics are now affine (and not simply linear) in the inputs.

Finally, in sections 5.6 and 5.7, we discuss the same sensor fusion problem, in 3D. This last example is very important not only in technological sciences but also in neuroscience. As it will be seen, our analysis will allow us to obtain compelling results about the problem of visual-vestibular integration for self-motion perception in mammals.

Note that the method is very powerful and can be used to automatically obtain the observability properties of any system that satisfies (2.1), i.e., independently of the state dimension (intuitive reasoning becomes often prohibitive for systems characterized by high-dimensional states), independently of the type of nonlinearity and in general independently of the system complexity. To this regard, note that the method can be implemented automatically, by using a simple symbolic computation tool (in our examples, in most of cases we executed the computation manually, and, in the hardest cases, we simply used the toolbox of MATLAB).

5.1 Unicycle with one input unknown

5.1.1 The system

We consider a vehicle that moves on a $2D$-environment. The configuration of the vehicle in a global reference frame, can be characterized through the vector $[x_v,\ y_v,\ \theta]^T$ where x_v and y_v are the Cartesian vehicle coordinates, and θ is the vehicle orientation. We assume that the dynamics of this vector satisfy the unicycle differential equations:

$$\begin{bmatrix} \dot{x}_v & = v\cos\theta \\ \dot{y}_v & = v\sin\theta \\ \dot{\theta} & = \omega \end{bmatrix} \tag{5.1}$$

where v and ω are the linear and the rotational vehicle speed, respectively, and they are the system inputs. We consider the following three cases of output (see also figure 5.1 for an illustration):

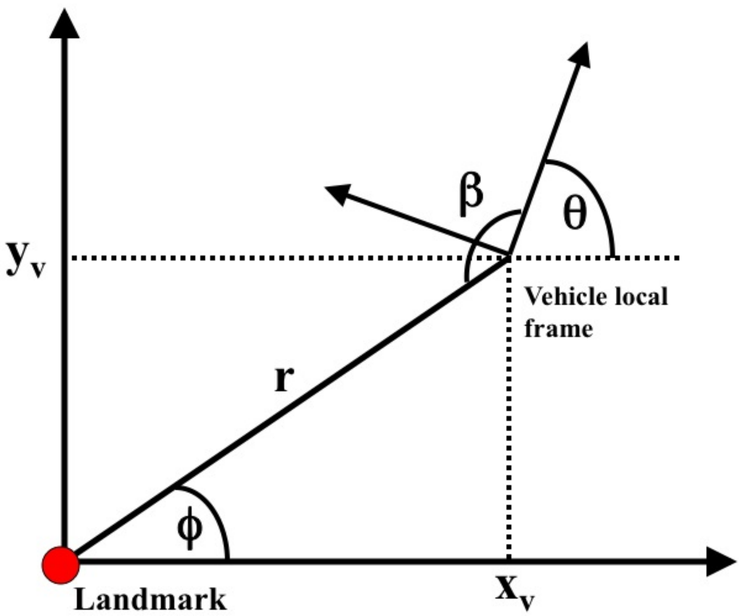

Figure 5.1: The vehicle state in cartesian and polar coordinates ($[x_v, y_v, \theta]^T$ and $[r, \phi, \theta]^T$, respectively) together with the three considered outputs (r, β and ϕ).

1. the distance (r) from the origin (e.g., a landmark is at the origin and its distance is measured by a range sensor);
2. the bearing angle (β) of the origin in the local frame (e.g., a landmark is at the origin and its bearing angle is measured by an on-board camera);
3. the bearing angle (ϕ) of the vehicle in the global frame (e.g., a camera is placed at the origin).

Case	Dimension of Ω
1^{st}: $y = r$, $u = \omega$, $w = v$	2
2^{nd}: $y = r$, $w = \omega$, $u = v$	2
3^{d}: $y = \theta - \phi$, $u = \omega$, $w = v$	1
4^{th}: $y = \theta - \phi$, $w = \omega$, $u = v$	2
5^{th}: $y = \phi$, $u = \omega$, $w = v$	2
6^{th}: $y = \phi$, $w = \omega$, $u = v$	3

Table 5.1: Dimension of the observable codistribution (Ω) obtained by following intuitive reasoning

We can analytically express the output in terms of the state. We remark that the expressions become very simple if we adopt polar coordinates: $r \triangleq \sqrt{x_v^2 + y_v^2}$, $\phi = atan\frac{y_v}{x_v}$. We have, for the three cases, $y = r$, $y = \pi - (\theta - \phi)$ and $y = \phi$, respectively. For each of these three cases, we consider the following two cases: v is known, ω is unknown; v is unknown, ω is known. Hence, we have six cases. The dynamics of the unicycle in polar coordinates become:

$$\left[\begin{array}{l} \dot{r} = v\cos(\theta - \phi) \\ \dot{\phi} = \dfrac{v}{r}\sin(\theta - \phi) \\ \dot{\theta} = \omega \end{array}\right. \tag{5.2}$$

5.1.2 Intuitive procedure to obtain the observability properties

By using the observability rank condition in [18], we easily obtain that, when both the inputs are known, the dimension of the observable codistribution is 2 for the first two observations ($y = r$ and $y = \theta - \phi$) and 3 for the last one ($y = \phi$). In particular, for the first two observations all the initial states rotated around the vertical axis are indistinguishable. When one of the inputs misses, this unobservable degree of freedom obviously remains. On the other hand, when the linear speed is unknown (i.e., it acts as an unknown input ($w = v$)) and the observation is an angle (second and third observation, i.e., $y = \theta - \phi$ and $y = \phi$, respectively), we lose a further degree of freedom, which corresponds to the absolute scale. In table 5.1 we provide the dimension of the observable codistribuion obtained by following this intuitive reasoning for the six considered cases.

5.1.3 Analytic results

We now derive the observability properties by applying the analytic criterion described in chapter 4. For all the cases we have $m_u = 1$. Hence, we adopt the following notation: $f \triangleq f^1$ (for the vector field in (2.2)) and $\phi_m \triangleq {}^1\phi_m$ (for the vectors defined by algorithm 2). We consider the six cases defined in section 5.1.1, separately.

First Case: $y = r$, $u = \omega$, $w = v$

We have:

$$f=\begin{bmatrix}0\\0\\1\end{bmatrix}\quad g=\begin{bmatrix}\cos(\theta-\phi)\\\frac{\sin(\theta-\phi)}{r}\\0\end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g=\cos(\theta-\phi)$, which does not vanish, in general.

Step 2

We have: $\Omega_0=span\{[1,0,0]\}$. Additionally, $\Omega_1=\Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0}L^1_g=\mathcal{L}_fL^1_g=-\sin(\theta-\phi)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau\triangleq\frac{L^2_g}{(L^1_g)^2}=\frac{\tan^2(\theta-\phi)}{r}$ and

$$\mathcal{D}\tau=\frac{\tan(\theta-\phi)}{r}\left[-\frac{\tan(\theta-\phi)}{r},-\frac{2}{\cos^2(\theta-\phi)},\frac{2}{\cos^2(\theta-\phi)}\right]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute ϕ_1. We obtain:

$$\phi_1=\begin{bmatrix}-\tan(\theta-\phi)\\\frac{1}{r}\\0\end{bmatrix}$$

and

$$\Omega_2=span\left\{[1,0,0],\ \left[0,\frac{1}{\cos^2(\theta-\phi)},-\frac{1}{\cos^2(\theta-\phi)}\right]\right\}$$

It is immediate to check that $\mathcal{D}\tau\in\Omega_2$, meaning that $m'=2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3=\Omega_2$ meaning that $m^*=2$ and $\Omega^*=\Omega_2$

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 5.1 (second line).

Second Case: $y = r$, $u = v$, $w = \omega$

We have:

$$f = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We easily obtain $\mathcal{L}_g h = 0$. We consider the function $\mathcal{L}_f h \in \mathcal{F}$. We have: $\mathcal{L}_f h = \cos(\theta - \phi)$ and $\mathcal{L}_g \mathcal{L}_f h = -\sin(\theta - \phi)$, which does not vanish, in general. Hence, we can proceed with the steps in chapter 4 by setting:

$$h = \cos(\theta - \phi)$$

We obtain: $L_g^1 = -\sin(\theta - \phi)$

Step 2

We have:

$$\Omega_0 = span\{[1, 0, 0], \ [0, \sin(\theta - \phi), \ -\sin(\theta - \phi)]\}$$

as long as the function r is also a system output. Additionally, $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L_g^1 = \mathcal{L}_f L_g^1 = \frac{\sin(\theta-\phi)\cos(\theta-\phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have:

$$\tau = -\frac{\cos(\theta - \phi)}{\sin^2(\theta - \phi)}$$

Step 5

By a direct computation we obtain $\Omega_2 = \Omega_1$. Additionally, it is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 5.1 (third line).

Third Case: $y = \theta - \phi$, $u = \omega$, $w = v$

We have:

$$f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad g = \begin{bmatrix} \cos(\theta-\phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g = -\frac{\sin(\theta-\phi)}{r}$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[0, -1, 1]\}$ and $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L^1_g = \mathcal{L}_f L^1_g = -\frac{\cos(\theta-\phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = 2\cot(\theta - \phi)$ and

$$\mathcal{D}\tau = \frac{2}{\sin^2(\theta-\phi)} [0, 1, -1]$$

Step 5

By a direct computation we obtain $\Omega_2 = \Omega_1$. Additionally, it is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

The dimension of the observable codistribution is 1. We conclude that the state is not weakly locally observable. This result agrees with the one in table 5.1 (fourth line). Note that, the new unobservable direction with respect to the case when both inputs are known, is precisely the absolute scale, since the vector $\mathcal{D}r = [1, 0, 0] \notin \Omega^*$.

Fourth Case: $y = \theta - \phi$, $u = v$, $w = \omega$

We have:

$$f = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g = 1 \neq 0$.

Step 2

By a direct computation we obtain: $\Omega_0 = span\{[0, -1, 1]\}$ and

$$\Omega_1 = span\left\{[0, -1, 1], \left[-\frac{\sin(\theta - \phi)}{r^2}, -\frac{\cos(\theta - \phi)}{r}, \frac{\cos(\theta - \phi)}{r}\right]\right\}$$

Step 3

Since $L^1_g = 1$, it is immediate to realize that $\mathcal{L}_{\phi_j} L^1_g = 0$, for any integer $j \geq 0$. Hence, the considered system is in the special case considered by lemma 6 and we skip to step 6 by setting $m' = 0$.

Step 6

By a direct computation, it is possible to check that $\Omega_2 = \Omega_1$ meaning that $m^* = 1$ and $\Omega^* = \Omega_1$.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 5.1 (fifth line).

Fifth Case: $y = \phi$, $u = \omega$, $w = v$

We have:

$$f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad g = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g = \frac{\sin(\theta-\phi)}{r}$, which does not vanish, in general.

Step 2

We easily obtain $\Omega_0 = span\{[0, 1, 0]\}$ and $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L^1_g = \mathcal{L}_f L^1_g = \frac{\cos(\theta-\phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -2\cot(\theta-\phi)$ and

$$\mathcal{D}\tau = \frac{2}{\sin^2(\theta-\phi)}\,[0,-1,1]$$

Step 5

To compute Ω_2 we need to compute ϕ_1. We obtain:

$$\phi_1 = \begin{bmatrix} -r \\ \cot(\theta-\phi) \\ 0 \end{bmatrix}$$

and

$$\Omega_2 = span\left\{[0,1,0],\ \frac{1}{\sin^2(\theta-\phi)}\,[0,1,-1]\right\}$$

It is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation we obtain $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$, whose dimension is 2.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 5.1 (sixth line). Note that, the new unobservable direction with respect to the case when both inputs are known, is precisely the absolute scale, since the vector $\mathcal{D}r = [1,0,0] \notin \Omega^*$.

Sixth Case: $y = \phi$, $u = v$, $w = \omega$

We have:

$$f = \begin{bmatrix} \cos(\theta-\phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We easily obtain $\mathcal{L}_g h = 0$. We consider the function $\mathcal{L}_f h \in \mathcal{F}$. We have: $\mathcal{L}_f h = \frac{\sin(\theta-\phi)}{r}$ and $\mathcal{L}_g \mathcal{L}_f h = \frac{\cos(\theta-\phi)}{r}$, which does not vanish, in general. Hence, we can proceed with the steps in chapter 4 by setting:

$$h = \frac{\sin(\theta-\phi)}{r} \qquad L_g^1 = \frac{\cos(\theta-\phi)}{r}$$

Step 2

We have:

$$\Omega_0 = span\left\{[0,1,0], \ \left[-\frac{\sin(\theta-\phi)}{r^2}, -\frac{\cos(\theta-\phi)}{r}, \frac{\cos(\theta-\phi)}{r}\right]\right\}$$

as long as the function ϕ is also a system output. We compute Ω_1. By a direct computation, we obtain that its dimension is 3. Hence, we do not need to proceed with the remaining steps since we can directly conclude that the entire state is weakly locally observable. This result agrees with the one in table 5.1 (seventh line).

5.2 Unicycle in presence of an external disturbance

5.2.1 The system

We consider the same vehicle considered in section 5.1 and we adopt the same state $[x_v, \ y_v, \ \theta]^T$ to characterize its position and orientation. We assume that the vehicle motion is also affected by an unknown input that produces an additional (and unknown) robot speed (denoted by w) along a fixed direction (denoted by γ). Hence, the dynamics are characterized by the following differential equations:

$$\begin{bmatrix} \dot{x}_v & = v\cos\theta + w\cos\gamma \\ \dot{y}_v & = v\sin\theta + w\sin\gamma \\ \dot{\theta} & = \omega \end{bmatrix} \tag{5.3}$$

where v and ω are the linear and the rotational speed, respectively, in absence of the unknown input. We assume that these two speeds are known (we refer to them as to the known inputs), w is unknown (we refer to it as to the unknown input or disturbance) and γ is constant in time. See also figure 5.2 for an illustration.

We consider the same three cases of output considered in section 5.1. Additionally, we deal with both the cases when γ is known and unknown (in section 5.2.2 and 5.2.3, respectively).

5.2.2 Observability properties when the disturbance direction is known

The state is $[x_v, \ y_v, \ \theta]^T$ and its dynamics are provided by the three equations in (5.3), where γ is a known parameter. These equations are a special case of (2.2). From (5.3) and (2.2) we easily obtain: $m_u = 2$, $u_1 = v$, $u_2 = \omega$,

$$f^1 = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \quad f^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad g = \begin{bmatrix} \cos\gamma \\ \sin\gamma \\ 0 \end{bmatrix}$$

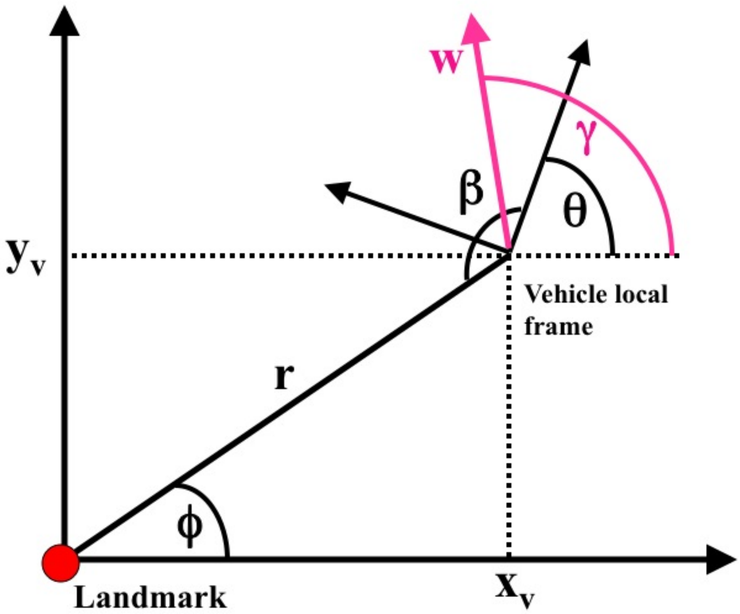

Figure 5.2: The vehicle state together with the three considered outputs.

We consider the three outputs separately. For the simplicity sake, we actually consider the following three outputs: $y = r^2 = x_v^2 + y_v^2$ instead of $y = r$, $y = \tan\beta = \frac{y_v - x_v \tan\theta}{x_v + y_v \tan\theta}$ instead of $y = \beta$ and $y = \tan\phi = \frac{y_v}{x_v}$ instead of $y = \phi$. Obviously, the result of the observability analysis does not change.

First Case: $y = r^2$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L_g^1 = 2(x_v \cos\gamma + y_v \sin\gamma)$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[x_v, y_v, 0]\}$ and $\Omega_1 = span\{[x_v, y_v, 0], [\cos\theta, \sin\theta, y_v \cos\theta - x_v \sin\theta]\ \}$.

Step 3

We have $\mathcal{L}_{{}^1\phi_0} L_g^1 = \mathcal{L}_{f^1} L_g^1 = 2\cos(\gamma - \theta)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{1}{2(x_v \cos\gamma + y_v \sin\gamma)^2}$ and

$$\mathcal{D}\tau = -\frac{1}{(x_v \cos\gamma + y_v \sin\gamma)^3} [\cos\gamma, \sin\gamma, 0]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute ${}^1\phi_1$ and ${}^2\phi_1$ through algorithm 2. We obtain: ${}^1\phi_1 = {}^2\phi_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. On the other hand, we obtain that $\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{D}\mathcal{L}_{f^1} h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Second Case: $y = \tan\beta$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2 \cos^2\theta + 2\sin\theta\cos\theta x_v y_v - y_v^2 \cos^2\theta + y_v^2}$, which does not vanish, in general.

Step 2

By an explicit computation (by using algorithm 3) we obtain that the dimension of Ω_0 is 1 and the dimension of Ω_1 is 2.

Step 3

We have $\mathcal{L}_{{}^2\phi_0} L^1_g = \mathcal{L}_{f^2} L^1_g \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{x_v \cos\gamma + y_v \sin\gamma + x_v \cos(\gamma - 2\theta) - y_v \sin(\gamma - 2\theta)}{y_v \cos\gamma - x_v \sin\gamma}$

Step 5

We need to compute Ω_2. Also in this case, we obtain that $\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{D}\mathcal{L}_{f^1} h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Third Case: $y = \tan\phi$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We have $L^1_g = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2}$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[-y_v, x_v, 0]\}$. In addition, by an explicit computation (by using algorithm 3) we obtain that the dimension of Ω_1 is 2.

Step 3

We have $\mathcal{L}_{^2\phi_0} L^1_g = \mathcal{L}_{f^2} L^1_g \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -\frac{2x_v^4 \cos\gamma}{x_v^4 \sin\gamma - x_v^3 y_v \cos\gamma}$

Step 5

We need to compute Ω_2. Also in this case, we obtain that $\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{D}\mathcal{L}_{f^1} h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

5.2.3 Observability properties when the disturbance direction is unknown

The state is $[x_v,\ y_v,\ \theta,\ \gamma]^T$ and its dynamics are provided by the following four equations:

$$\begin{bmatrix} \dot{x}_v & = v\cos\theta + w\cos\gamma \\ \dot{y}_v & = v\sin\theta + w\sin\gamma \\ \dot{\theta} & = \omega \\ \dot{\gamma} & = 0 \end{bmatrix} \tag{5.4}$$

From (5.4) and (2.2) we easily obtain: $m_u = 2$, $u_1 = v$, $u_2 = \omega$,

$$f^1 = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \\ 0 \end{bmatrix}, \quad f^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} \cos\gamma \\ \sin\gamma \\ 0 \\ 0 \end{bmatrix}$$

We consider the three outputs separately. As in section 5.2.2, we actually consider the following three outputs: $y = r^2 = x_v^2 + y_v^2$ instead of $y = r$, $y = \tan\beta = \frac{y_v - x_v \tan\theta}{x_v + y_v \tan\theta}$ instead of $y = \beta$ and $y = \tan\phi = \frac{y_v}{x_v}$ instead of $y = \phi$. Obviously, the result of the observability analysis does not change.

First Case: $y = r^2$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = r^2$ of section 5.2.2, i.e., $L^1_g = 2(x_v \cos\gamma + y_v \sin\gamma)$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[x_v, y_v, 0, 0]\}$ and $\Omega_1 = span\{[x_v, y_v, 0, 0], [\cos\theta, \sin\theta, y_v \cos\theta - x_v \sin\theta, 0]\ \}$.

Step 3

We have $\mathcal{L}_{^1\phi_0} L^1_g = \mathcal{L}_{f^1} L^1_g = 2\cos(\gamma - \theta)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{1}{2(x_v \cos\gamma + y_v \sin\gamma)^2}$, as in the case $y = r^2$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ, namely:

$$\mathcal{D}\tau = \frac{-1}{(x_v \cos\gamma + y_v \sin\gamma)^3} [\cos\gamma,\ \ \sin\gamma,\ \ 0,\ \ y_v \cos\gamma - x_v \sin\gamma]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute $^1\phi_1$ and $^2\phi_1$ through algorithm 2. We obtain: $^1\phi_1 =^2 \phi_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. By using algorithm 3 we compute Ω_2 and we obtain that its dimension is 3. Additionally, it is possible to verify that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

We conclude that the dimension of the observable codistribution is equal to $3 (< 4)$ and the state is not weakly locally observable. In particular, since the differential of every state component does not belong to Ω_2, we conclude that no state component is observable.

Second Case: $y = \tan\beta$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = \tan\beta$ of section 5.2.2, i.e., $L^1_g = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2 \cos^2\theta + 2\sin\theta\cos\theta x_v y_v - y_v^2 \cos^2\theta + y_v^2}$, which does not vanish, in general.

Step 2

We compute Ω_0 and Ω_1: their dimension are 1 and 2, respectively.

Step 3

We have $\mathcal{L}_{^2\phi_0} L^1_g = \mathcal{L}_{f^2} L^1_g \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{x_v \cos\gamma + y_v \sin\gamma + x_v \cos(\gamma - 2\theta) - y_v \sin(\gamma - 2\theta)}{y_v \cos\gamma - x_v \sin\gamma}$, as in the case $y = \tan\beta$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ.

Step 5

By using algorithm 3 we compute Ω_2 and we obtain that its dimension is 3. Additionally, it is possible to verify that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

We conclude that the dimension of the observable codistribution is equal to $3 (< 4)$ and the state is not weakly locally observable. In particular, since the differential of every state component does not belong to Ω_2, we conclude that no state component is observable.

Third Case: $y = \tan\phi$

We apply the analytic criterion in chapter 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = \tan\phi$ of section 5.2.2, i.e., $L^1_g = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2}$, which does not vanish, in general.

Step 2

We compute Ω_0 and Ω_1: their dimension are 1 and 2, respectively.

Step 3

We have $\mathcal{L}_{^2\phi_0} L^1_g = \mathcal{L}_{f^2} L^1_g \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -\frac{2x_v^4 \cos\gamma}{x_v^4 \sin\gamma - x_v^3 y_v \cos\gamma}$, as in the case $y = \tan\phi$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ.

γ	Output	State observability
known	$y = r$	yes
known	$y = \beta$	yes
known	$y = \phi$	yes
unknown	$y = r$	no
unknown	$y = \beta$	no
unknown	$y = \phi$	yes

Table 5.2: Weak local observability of the state in all the considered scenarios

Step 5

By using algorithm 3 we compute Ω_2 and we obtain that its dimension is 4. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Table 5.2 summarizes the results of the observability analysis carried out in this section. The reader can find in [43] the results of extensive simulations for the system studied in this section. These results clearly validate the analytic results of the previous observability analysis. In addition, we remark that the analytic results provided by our observability analysis are also understandable by following intuitive reasoning. By using the observability rank condition in [18], we easily obtain that, in absence of the unknown input, the dimension of the observable codistribution is 2 for the first two observations ($y = r$ and $y = \beta$) and 3 for the last one ($y = \phi$). In particular, for the first two observations, all the initial states rotated around the vertical axis are indistinguishable. In other words, in these two cases, the system exhibits a continuous symmetry [37]. In presence of the unknown input, when γ is known, the aforementioned system invariance is broken and the entire state becomes observable. When γ is unknown, the symmetry still remains (and obviously also concerns the new state component γ).

We conclude by remarking a very important aspect. The presence of an unknown input improves the observability properties of a system (this regards the case when γ is known). In particular, if $w = 0$ (absence of unknown input), the state becomes unobservable despite the knowledge of the unknown input (we know that it is zero), while, when $w \neq 0$, the state is observable even if w is unknown. Note that having an unknown input equal to zero is an event that occurs with zero probability and our theory accounts this fact since it is based on definition 3. To this regard, note also that the validity of theorem 1, which allows us to introduce the algorithms 2 and 3, holds when the unknown input is different from 0.

5.3 Vehicle moving in 3D in presence of a disturbance

5.3.1 The system

We consider a vehicle that moves in a $3D-$environment. We assume that the dynamics of the vehicle are affected by the presence of a disturbance (e.g., this could be an aerial vehicle in presence of wind). We assume that the direction of the disturbance is constant in time and a priori known. Conversely, the disturbance magnitude is unknown and time dependent. The vehicle is equipped with speed sensors (e.g., airspeed sensors in the case of an aerial vehicle), gyroscopes and a bearing sensor (e.g., monocular camera). We assume that all the sensors share

the same frame (in other words, they are extrinsically calibrated). Without loss of generality, we define the vehicle local frame as this common frame. The airspeed sensors measure the vehicle speed with respect to the air in the local frame. The gyroscopes provide the angular speed in the local frame. Finally, the bearing sensor provides the bearing angles of the features in the environment expressed in its own local frame. We consider the extreme case of a single point feature and, without loss of generality, we set the origin of the global frame at this point feature (see figure 5.3 for an illustration). Additionally, we assume that the $z-$axis of the global frame is aligned with the direction of the disturbance.

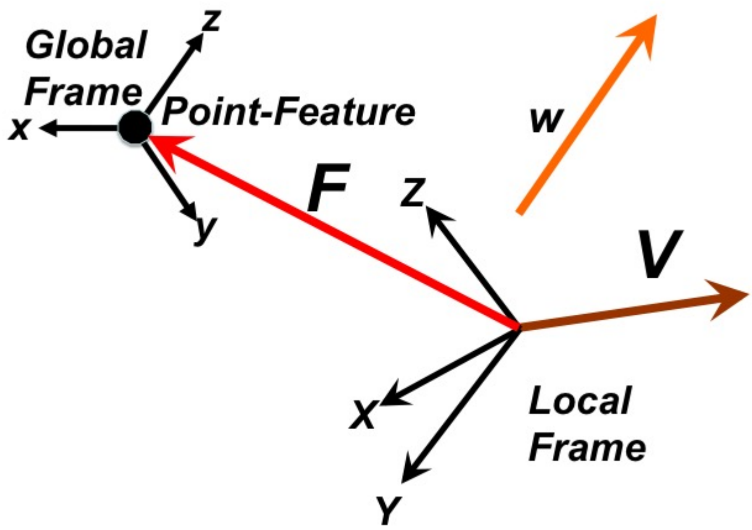

Figure 5.3: Local and global frame for the considered problem. The $z-$axis of the latter is aligned with the direction of the disturbance (assumed to be known and constant in time). The speed V is the vehicle speed with respect to the air, which differs from the ground speed because of the disturbance (w).

Our system can be characterized by the following state:

$$X \triangleq [x,\ y,\ z,\ q_t,\ q_x,\ q_y,\ q_z]^T \tag{5.5}$$

where $r = [x,\ y,\ z]$ is the position of the vehicle in the global frame and $q = q_t + q_x i + q_y j + q_z k$ is the unit quaternion that describes the transformation change between the global and the local frame. The dynamics are affected by the presence of the disturbance. The disturbance is characterized by the following vector (in the global frame):

$$\bar{w} = w \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{5.6}$$

where w is its unknown magnitude.

In the sequel, for each vector defined in the $3D$ space, the pedix q will be adopted to denote the corresponding imaginary quaternion. For instance, regarding the vehicle position, we have: $r_q = 0 + x\ i + y\ j + z\ k$. Additionally, we denote by V and Ω the following physical quantities:

- $V = [V_x,\ V_y,\ V_z]$ is the vehicle speed with respect to the air expressed in the local frame (hence, $w\ k + qV_q q^*$ is the vehicle speed with respect to the ground expressed in the global frame).
- $\Omega \triangleq [\Omega_x\ \Omega_y\ \Omega_z]$ is the angular speed (and $\Omega_q = 0 + \Omega_x\ i + \Omega_y\ j + \Omega_z\ k$).

The dynamics of the state are:

$$\begin{bmatrix} \dot{r}_q & = w\ k + qV_q q^* \\ \dot{q} & = \frac{1}{2} q\Omega_q \end{bmatrix} \tag{5.7}$$

The monocular camera provides the position of the feature in the local frame ($F_q = -q^* r_q q$) up to a scale. Hence, it provides the ratios of the components of F:

$$h_{cam}(X) \triangleq [h_u,\ h_v]^T = \left[\frac{(q^* r_q q)_x}{(q^* r_q q)_z},\ \frac{(q^* r_q q)_y}{(q^* r_q q)_z}\right]^T \tag{5.8}$$

where the pedices x, y and z indicate respectively the i, j and k component of the corresponding quaternion. We have also to consider the constraint $q^* q = 1$. This provides the further observation:

$$h_{const}(X) \triangleq q^* q \tag{5.9}$$

Our system is characterized by the state in (5.5), the dynamics in (5.7) and the three outputs h_u, h_v and h_{const} in (5.8) and (5.9).

5.3.2 Observability in absence of disturbance

Our system is characterized by the state in (5.5), the dynamics in (5.7) with $w = 0$ and the three outputs h_u, h_v and h_{const} in (5.8) and (5.9).

By comparing (5.7) with (2.2) we obtain that our system is characterized by six known inputs ($m_u = 6$) that are: $u_1 = \Omega_x$, $u_2 = \Omega_y$, $u_3 = \Omega_z$, $u_4 = V_x$, $u_5 = V_y$ and $u_6 = V_z$. Additionally, we obtain:

$$f^1 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_x \\ q_t \\ q_z \\ -q_y \end{bmatrix},\ f^2 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_y \\ -q_z \\ q_t \\ q_x \end{bmatrix},\ f^3 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_z \\ q_y \\ -q_x \\ q_t \end{bmatrix},$$

$$f^4 = \begin{bmatrix} q_t^2 + q_x^2 - q_y^2 - q_z^2 \\ 2q_t q_z + 2q_x q_y \\ 2q_x q_z - 2q_t q_y \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ f^5 = \begin{bmatrix} 2q_x q_y - 2q_t q_z \\ q_t^2 - q_x^2 + q_y^2 - q_z^2 \\ 2q_t q_x + 2q_y q_z \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$f^6 = \begin{bmatrix} 2q_tq_y + 2q_xq_z \\ 2q_yq_z - 2q_tq_x \\ q_t^2 - q_x^2 - q_y^2 + q_z^2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Finally, in absence of disturbance we have:

$$g = [0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T$$

In this case we can apply the observability rank condition, i.e., algorithm 1, to obtain the observable codistribution. We compute the codistribution Ω_0 by computing the differentials of the three functions h_u, h_v and h_{const}. We obtain that this codistribution has dimension equal to 3. We use algorithm 1 to compute Ω_1. We obtain that its dimension is 4. In particular, the additional covector is obtained by the differential of the following Lie derivative:

$$\mathcal{L}_{f^4} h_u$$

In other words:

$$\Omega_1 = span\left\{\mathcal{D}h_u,\ \mathcal{D}h_v,\ \mathcal{D}h_{const},\ \mathcal{D}\mathcal{L}_{f^4}h_u,\right\}$$

All the remaining first order Lie derivatives have differential that is in the above codistribution. Additionally, by an explicit computation, it is easy to realize that $\Omega_2 = \Omega_1$. This means that algorithm 1 has converged and the observable codistribution is Ω_1.
By an explicit computation, it is possible to check that the differentials of the components of the vector F belong to Ω_1. This means that all the observable modes are the components of F, i.e., the position of the feature in the local frame (obviously, the fourth observable mode is the norm of the quaternion). In particular, no component of the vehicle orientation is observable.

5.3.3 Observability in presence of the disturbance

We now consider the case when the dynamics are affected by the presence of the disturbance. By comparing (5.7) with (2.2) we obtain that the vector fields that characterize the dynamics are the same that characterize the dynamics in absence of disturbance with the exception of the last one, which becomes:

$$g = [0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0]^T$$

To derive the observability properties we apply the proposed analytic tool, by following the seven steps provided in chapter 4.

First Step

We start by computing the Lie derivatives of the outputs h_u, h_v and h_{const} along the vector field g. We find that the result differs from zero for the first two outputs. Hence, we use the first output (h_u) to define L_g^1 (we could choose also the second output h_v). In particular, we obtain: $L_g^1 \triangleq \mathcal{L}_g h_u =$

$$\frac{-y(2q_tq_z - 2q_xq_y) - x(q_t^2 - q_x^2 + q_y^2 - q_z^2)}{[z(q_t^2 - q_x^2 - q_y^2 + q_z^2) + 2x(q_tq_y + q_xq_z) + 2y(q_yq_z - q_tq_x)]^2}$$

Second Step

We compute the codistribution Ω_0 by computing the differentials of the three functions h_u, h_v and h_{const}. This coincides with the case without disturbance, and we obtain that this codistribution has dimension equal to 3.

We use algorithm 3 to compute Ω_1. We obtain that its dimension is 5. In particular, the additional two independent covectors are obtained by the differentials of the following two Lie derivatives:

$$\mathcal{L}_{f^4} h_u, \quad \mathcal{L}_{\frac{g}{L_g^1}} h_v$$

In other words:

$$\Omega_1 = span\left\{\mathcal{D}h_u,\ \mathcal{D}h_v,\ \mathcal{D}h_{const},\ \mathcal{D}\mathcal{L}_{f^4} h_u,\ \mathcal{D}\mathcal{L}_{\frac{g}{L_g^1}} h_v\right\}$$

All the remaining first order Lie derivatives have differential that is in the above codistribution.

Third Step

We compute ${}^1\phi_1$, ${}^2\phi_1$, ${}^3\phi_1$, ${}^4\phi_1$, ${}^5\phi_1$ and ${}^6\phi_1$ by using algorithm 2. We obtain that all these vectors vanish. As a result, all the subsequent steps of algorithm 2 provide null vectors. Therefore, the assumptions of lemma 6 are trivially met. We set $m' = 0$ and we skip to the sixth step.

Sixth Step

We use algorithm 3 to compute Ω_2 and we obtain:

$$\Omega_2 = \Omega_1 + span\left\{\ \mathcal{D}\mathcal{L}_{f^4}\mathcal{L}_{\frac{g}{L_g^1}} h_v\right\}$$

Hence, its dimension is 6. Finally, by using again algorithm 3 it is possible to compute Ω_3 and to check that $\Omega_3 = \Omega_2$. This means that the algorithm has converged and the observable codistribution is $\Omega^* = \Omega_2$.

Seventh Step

By computing the distribution orthogonal to the codistribution Ω^* we can find the continuous symmetry that characterizes the unobservable space [37]. By an explicit computation we obtain the following vector:

$$\left[-y,\ x,\ 0,\ -\frac{q_z}{2},\ -\frac{q_y}{2},\ \frac{q_x}{2},\frac{q_t}{2}\right]^T$$

This symmetry corresponds to an invariance with respect to a rotation around the $z-$axis of the global frame. This means that we have a single unobservable mode that is the yaw in the global frame[1]. We conclude by remarking that the presence of the disturbance, even if its magnitude is unknown and is not constant, makes observable the roll and the pitch angles. This result is similar to the result that we obtain in the case of visual and inertial sensor fusion in presence of gravity. The presence of gravity makes observable the roll and the pitch angles, even if its magnitude is unknown [38, 39]. What it is non intuitive in the case now investigated, is that, not only the magnitude of the disturbance is unknown, but it is also time dependent.

The reader can find in [43] the results of simulations for the system studied in this section. These results clearly validate the analytic results of the previous observability analysis.

[1]Note that the chosen global frame is aligned with he direction of the disturbance (fig. 5.3). Hence, what is unobservable is a rotation around the direction of the disturbance.

5.4 Visual-inertial sensor fusion: the planar case with calibrated sensors

We consider a sensor suit that consists of inertial sensors (IMU) and a bearing sensor (camera). We assume that this sensor suit moves on a $2D-$environment. Without loss of generality, we define the sensor suit local frame as the IMU frame. The inertial sensors measure the sensor suit acceleration and the angular speed. In $2D$, the acceleration is a two dimensional vector and the angular speed is a scalar. The camera provides the bearing angle of the features in its own local frame.

In this section, we derive the observability properties in the case when the IMU only provides the acceleration along a single axis (instead of two). In other words, we will consider the other component of the acceleration and the angular speed as unknown inputs. Therefore, with respect to the examples investigated in sections 5.1 and 5.2, the system is characterized by two independent unknown inputs and, as we will see, the dynamics are affine in the inputs and not simply linear (i.e., the dynamics are characterized by a non null vector g_0). Finally, we directly investigate the most challenging case of a single point feature.

In general, the camera frame does not coincide with the sensor suit frame. Additionally, the measurements provided by the IMU are in general biased. We start our investigation in the simple case when the camera frame coincides with the sensor suit frame and the inertial measurements are unbiased. Then, in section 5.5, we relax both these assumptions by also assuming that the camera is extrinsically uncalibrated (i.e., the transformation between the IMU frame and the camera frame is unknown).

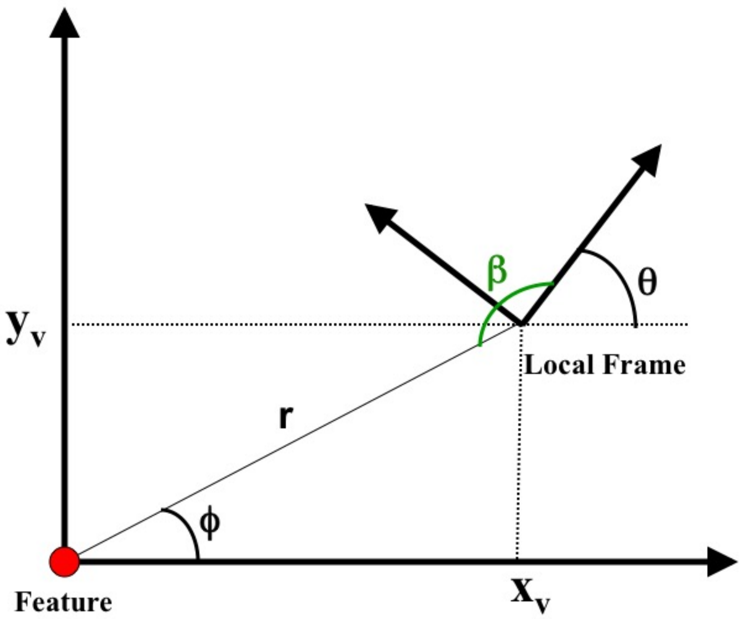

Figure 5.4: Visual-inertial sensor fusion in $2D$ with calibrated sensors and in the case of a single point feature. The global frame, the local frame and the bearing observation (β).

The state that characterizes our system is $[x_v, y_v, v_x, v_y, \theta]^T$, where the first two components are the sensor suit position, the second two components the sensor suit speed and the last component the sensor suit orientation. All these quantities are expressed in a common global frame (see fig 5.4 for an illustration). The dynamics are:

$$\left[\begin{array}{ll} \dot{x}_v & = v_x \\ \dot{y}_v & = v_y \\ \dot{v}_x & = \cos\theta A_x - \sin\theta A_y \\ \dot{v}_y & = \sin\theta A_x + \cos\theta A_y \\ \dot{\theta} & = \omega \end{array}\right. \tag{5.10}$$

where $[A_x, A_y]^T$ is the sensor suit acceleration expressed in the sensor suit frame and ω the angular speed. In order to have a simpler expression for the output, it is better to work in polar coordinates. In other words, we set:

- $r \triangleq \sqrt{x_v^2 + y_v^2}$;
- $\phi \triangleq \arctan\left(\frac{y_v}{x_v}\right)$;
- $v \triangleq \sqrt{v_x^2 + v_y^2}$;
- $\alpha \triangleq \arctan\left(\frac{v_y}{v_x}\right)$;

Hence, we define the state:

$$X = [r,\ \phi,\ v,\ \alpha,\ \theta]^T \tag{5.11}$$

Its dynamics are:

$$\left[\begin{array}{ll} \dot{r} & = v\cos(\alpha - \phi) \\ \dot{\phi} & = \frac{v}{r}\sin(\alpha - \phi) \\ \dot{v} & = A_x\cos(\alpha - \theta) + A_y\sin(\alpha - \theta) \\ \dot{\alpha} & = -\frac{A_x}{v}\sin(\alpha - \theta) + \frac{A_y}{v}\cos(\alpha - \theta) \\ \dot{\theta} & = \omega \end{array}\right. \tag{5.12}$$

Without loss of generality, we assume that the feature is positioned at the origin of the global frame. The camera provides the angle $\beta = \pi - \theta + \phi$. Hence, we can perform the observability analysis by using the output (we ignore π):

$$y = h(X) = \phi - \theta \tag{5.13}$$

We consider the system characterized by the state in (5.11), the dynamics in (5.12) and the output in (5.13) under the assumptions that the IMU only provides the acceleration along a single axis. Without loss of generality, we assume that it provides A_x. Hence, we have that $u = A_x$ is a known input while $w_1 = \omega$ and $w_2 = A_y$ are two unknown inputs (or disturbances). By comparing (5.12) with (2.1) we obtain:

$$g^0 = \begin{bmatrix} v\cos(\alpha-\phi) \\ \frac{v}{r}\sin(\alpha-\phi) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad f \triangleq f^1 = \begin{bmatrix} 0 \\ 0 \\ \cos(\alpha-\theta) \\ -\frac{1}{v}\sin(\alpha-\theta) \\ 0 \end{bmatrix}$$

$$g^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad g^2 = \begin{bmatrix} 0 \\ 0 \\ \sin(\alpha-\theta) \\ \frac{1}{v}\cos(\alpha-\theta) \\ 0 \end{bmatrix}$$

Before proceeding with the steps summarized in section 4.2, we need to check that the system is in canonic form. In the case is not, we need to perform the system canonization, as explained in appendix A.2.

5.4.1 System Canonization

We follow the procedure described in appendix A.2. For, we start by building the space of functions $\mathcal{F}^0$ and the codistribution $\mathcal{DF}^0$. Since $\mathcal{L}_f h = 0$, we obtain that $\mathcal{F}^0$ only contains the function $h(X)$ in (5.13) and $\mathcal{DF}^0 = span\{[0, 1, 0, 0, -1]\}$. Additionally, the space $\mathcal{L}_G\mathcal{F}^0$ only contains $\mathcal{L}_G h = \frac{v\sin(\alpha-\phi)}{r} - w_1$. We easily obtain that $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^0 = span\{[-1, 0]\}$. Hence, its dimension is $1 < m_w = 2$ and the system is not in canonic form. Since the covectors in $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^0$ have the second entry equal to zero, we do not need to change the coordinates according to (A.4). In accordance with (A.6), we include w_1 in the state:

$$X \to [X^T \ w_1]^T$$

We proceed by computing $\mathcal{F}^1$. Since $\mathcal{L}_f^2\mathcal{L}_G h = 0$, $\mathcal{F}^1$ only contains the functions: h, $\mathcal{L}_G h$ and $\mathcal{L}_f\mathcal{L}_G h$. In particular, they are independent, i.e., the codistribution $\mathcal{DF}^1$ has dimension equal to 3. We compute the space $\mathcal{L}_G\mathcal{F}^1$. Then, we define the new unknown input vector in accordance with (A.7):

$$^1w = [w_1^{(1)} \ w_2]$$

We obtain $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^1$ by computing the differentials (with respect to ^1w) of the functions in $\mathcal{L}_G\mathcal{F}^1$. By a direct computation we obtain that the dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^1$ is $1 < m_w = 2$ and the system is not in canonic form. This time, the second entry of the covectors in $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^1$ is not automatically 0. Hence, we need to change the coordinates in accordance with (A.4). We obtain:

$$w_1^{(1)} \to w_1^{(1)} - \frac{\cos(\theta-\phi)}{r} w_2$$
$$w_2 \to w_2$$

In accordance with (A.6), we include $w_1^{(1)}$ in the state:

$$X \to [X^T \ w_1^{(1)}]^T$$

We proceed by computing $\mathcal{F}^2$. In this case we obtain that the independent functions in $\mathcal{F}^2$, i.e., the ones whose differentials with respect to X generate $\mathcal{DF}^2$, are: h, $\mathcal{L}_G h$, $\mathcal{L}_G^2 h$, $\mathcal{L}_f\mathcal{L}_G^2 h$

and $\mathcal{L}_f^2\mathcal{L}_G^2 h$. We compute the space $\mathcal{L}_G\mathcal{F}^2$. Then, we define the new unknown input vector in accordance with (A.7):

$$^2w = [w_1^{(2)}\ w_2]$$

We obtain $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^2$ by computing the differentials (with respect to ^2w) of the functions in $\mathcal{L}_G\mathcal{F}^2$. By a direct computation we obtain that the dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^2$ is $2 = m_w$ and the system is in canonic form. In particular, $\mathcal{D}\mathcal{F}^2$ is generated starting from the following two functions in $\mathcal{F}^2$: $\mathcal{L}_G^2 h$, $\mathcal{L}_f\mathcal{L}_G^2 h$.

5.4.2 Observability properties

In accordance with the derivation in 5.4.1, we have the canonic form of our system. We denote by $\bar{\omega}$ the angular acceleration. The state is:

$$X = [r,\ \phi,\ v,\ \alpha,\ \theta,\ \omega,\ \bar{\omega}]^T \tag{5.14}$$

The known input and the second unknown input remain the same ($u = A_x$, $w_2 = A_y$). The first unknown input becomes:

$$w_1 = \dot{\bar{\omega}}$$

Regarding the dynamics, the first five components of the state satisfy (5.12). The last two components satisfy the following equations: $\dot{\omega} = \bar{\omega}$ and $\dot{\bar{\omega}} = w_1$.

By comparing the new dynamics with (2.1) we obtain:

$$g^0 = \begin{bmatrix} v\cos(\alpha-\phi) \\ \frac{v}{r}\sin(\alpha-\phi) \\ 0 \\ 0 \\ \omega \\ \bar{\omega} \\ 0 \end{bmatrix}, \quad f \triangleq f^1 = \begin{bmatrix} 0 \\ 0 \\ \cos(\alpha-\theta) \\ -\frac{1}{v}\sin(\alpha-\theta) \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$g^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad g^2 = \begin{bmatrix} 0 \\ 0 \\ \sin(\alpha-\theta) \\ \frac{1}{v}\cos(\alpha-\theta) \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We are now ready to apply the method in 4.2.

First step

In accordance with the derivation in 5.4.1, we can select h_1 and h_2 as follows:

$$h_1 = \mathcal{L}_G^2 h = -\bar{\omega} - \frac{v^2\sin(2\alpha-2\phi)}{r^2}$$
$$h_2 = \mathcal{L}_f\mathcal{L}_G^2 h = -2\frac{v\sin(\alpha-2\phi+\theta)}{r^2}$$

From (3.3) we obtain:

$$
\begin{aligned}
\mu_1^1 =& \mathcal{L}_{g^1} h_1 = -1 \\
\mu_1^2 =& \mathcal{L}_{g^2} h_1 = -2v\frac{\cos(\alpha - 2\phi + \theta)}{r^2} \\
\mu_2^1 =& \mathcal{L}_{g^1} h_2 = 0 \\
\mu_2^2 =& \mathcal{L}_{g^2} h_2 = \frac{2 - 4\cos(\phi - \theta)^2}{r^2}
\end{aligned}
$$

which is non singular.

Second step

We compute the inverse of the previous tensor. We easily obtain:

$$
\begin{aligned}
\nu_1^1 =& -1 \\
\nu_1^2 =& v\frac{\cos(\alpha - 2\phi + \theta)}{\cos(2\phi - 2\theta)} \\
\nu_2^1 =& 0 \\
\nu_2^2 =& -\frac{r^2}{2\cos(2\phi - 2\theta)}
\end{aligned}
$$

Additionally, we obtain from (3.6):

$$
\begin{aligned}
\mu_1^0 =& \mathcal{L}_{g^0} h_1 = \frac{2v^3 \sin(3\alpha - 3\phi)}{r^3} \\
\mu_2^0 =& \mathcal{L}_{g^0} h_2 = 2v\frac{2v\sin(2\alpha - 3\phi + \theta) - r\omega\cos(\alpha - 2\phi + \theta)}{r^3}
\end{aligned}
$$

Finally, from (3.9) we obtain:

$$
\hat{g}^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad \hat{g}^2 = \frac{1}{2v\cos(2\phi - 2\theta)} \begin{bmatrix} 0 \\ 0 \\ -vr^2\sin(\alpha - \theta) \\ -r^2\cos(\alpha - \theta) \\ 0 \\ -vr\cos(\theta - \phi) \\ 2v^2\cos(\alpha - 2\phi + \theta) \end{bmatrix}
$$

We do not provide the expression of $\hat{g}^0$ for the sake of brevity (its expression is more complex).

Third step

We compute the differential of h_1 and h_2 (obtained at the first step) with respect to the state in (5.14). By a direct computation we obtain that the dimension of Ω_0 is 2. We compute Ω_1 by using the recursive step of algorithm 5 (we can ignore its last two terms, since the outputs are among the functions h_1 and h_2). The dimension of Ω_1 is 3.

Fourth step

We have $\mathcal{L}_{\phi_0}\mathcal{L}_{g^2}h_1 = \mathcal{L}_f\mu_1^2 = \frac{4\sin^2(\phi-\theta)-2}{r^2}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 14 and we need to continue with step 5.

Fifth step

We compute the three-index tensor in (3.11). We remind the reader that we can consider the lower index as a Latin index since the components of the tensor when this index is zero vanish. We obtain: $\mathcal{T}_1^{01} = \mathcal{T}_2^{01} = \mathcal{T}_1^{10} = \mathcal{T}_2^{10} = \mathcal{T}_1^{11} = \mathcal{T}_2^{11} = \mathcal{T}_1^{12} = \mathcal{T}_2^{12} = \mathcal{T}_1^{21} = \mathcal{T}_2^{21} = \mathcal{T}_2^{22} = 0$,

$$\mathcal{T}_2^{02} = \frac{2v\sin(2\alpha-5\phi+3\theta)+2v\sin(\phi-\theta)-r^3\omega\sin(2\theta-2\phi)+r^3\omega\sin(2\phi-2\theta)}{r^3(1-2\sin(\phi-\theta)^2)}$$

$$\mathcal{T}_1^{20} = \frac{v\sin(2\phi-2\theta)(2v\sin(2\alpha-3\phi+\theta)-r\omega\cos(\alpha-2\phi+\theta))}{r\cos(2\phi-2\theta)^2} - 3\frac{v^2\cos(3\phi-2\alpha-\theta)+v^2\cos(2\alpha-3\phi+\theta)}{2r\cos(2\phi-2\theta)}$$

$$\mathcal{T}_2^{20} = \frac{2r\omega\sin(2\phi-2\theta)-3v\cos(\alpha-3\phi+2\theta)+v\cos(\phi-\alpha)-4v\cos(3\phi-\alpha-2\theta)}{2r\cos(2\phi-2\theta)}$$

$\mathcal{T}_1^{22} = \frac{r^2\sin(2\phi-2\theta)}{2\sin(2\phi-2\theta)^2-2}$. We do not provide the expression of $\mathcal{T}_1^{00}$,$\mathcal{T}_2^{00}$ and $\mathcal{T}_1^{02}$ for the sake of brevity (their expression is more complex).

Sixth step

We need to compute Ω_2 and, in order to do this, we need to compute ${}^0\phi_1, {}^1\phi_1$ and ${}^2\phi_1$, by using algorithm 4. We obtain that only the first one differs from the null vector and it is:

$${}^0\phi_1 = \begin{bmatrix} -\cos(\theta-\phi) \\ -\frac{\sin(\theta-\phi)}{r} \\ -\omega\sin(\theta-\alpha) \\ \frac{\omega\cos(\theta-\alpha)}{v} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence, by using algorithm 5 we can compute Ω_2. Its dimension is 6. In addition, we obtain that the differentials of all the components of the tensor $\mathcal{T}_j^{\alpha i}$ belong to Ω_2, meaning that $m' = 2$.

Seventh step

By an explicit computation (by using the subsequent step of both algorithms 4 and 5), we obtain $\Omega_3 = \Omega_2$. Hence, algorithm 5 has coverged and $\Omega^* = \Omega_2$.

Eighth step

By computing the distribution orthogonal to the codistribution Ω_2 we can find the continuous symmetry that characterizes the unobservable space [37]. By an explicit computation we obtain the vector: $[0,\ 1,\ 0,\ 1,\ 1,\ 0,\ 0]^T$. This symmetry corresponds to an invariance with respect to a rotation around the vertical axis[2]. Note that, the absolute scale is invariant with respect to

[2] If instead of working in polar coordinates we adopted Cartesian coordinates, we obtained the following symmetry: $[y_v,\ -x_v,\ v_y,\ -v_x,\ 1,\ 0,\ 0]^T$.

a rotation around the vertical axis meaning that it can be estimated. This also holds for the sensor suit speed in the local frame.

We remark that, the invariance that corresponds to the continuous symmetry detected by our observability analysis, would also exist by having a complete IMU, i.e., when the IMU provides the acceleration along the two axes and the angular speed. This means that the information obtained by fusing the measurements from a camera with the measurements from a complete IMU is redundant.

5.5 Visual-inertial sensor fusion: the planar case with uncalibrated sensors

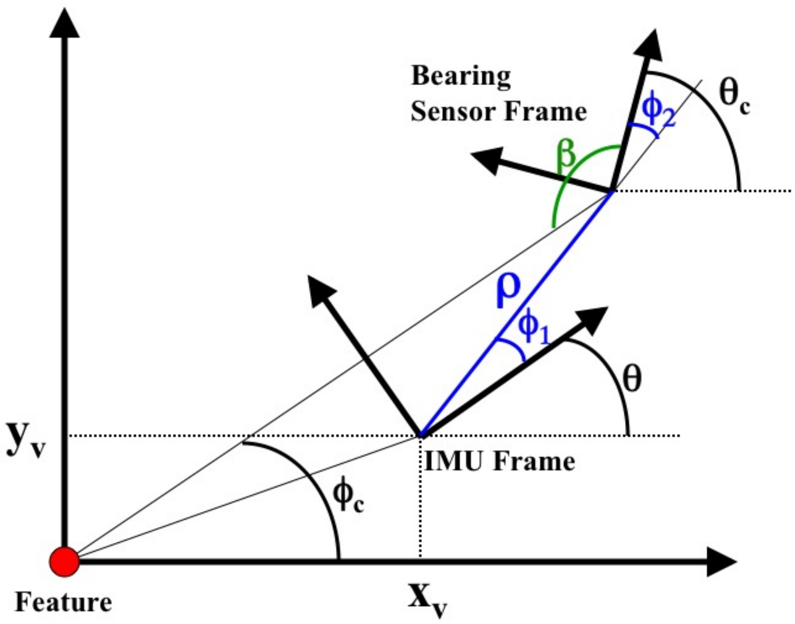

Figure 5.5: Visual-inertial sensor fusion in $2D$ with uncalibrated sensors and in the case of a single point feature. The global frame, the IMU frame and the bearing sensor frame are displayed together with the parameters that characterize the sensor suit.

We consider now the case when the measurements from the accelerometer are biased and the camera is extrinsically uncalibrated, i.e., the camera position and orientation in the sensor suit frame is unknown. We characterize the camera configuration in the sensor suit frame with the three parameters: ρ, ϕ_1 and ϕ_2 (see fig. 5.5 for an illustration). As in the previous case, we are assuming that the IMU only provides the acceleration along the X-axis of the local frame (A_x). Hence, we consider again the angular speed (ω) and the other component of the acceleration (A_y), as two independent unknown inputs. We characterize our system by the following state

(in this case polar coordinates are useless):

$$X = [x_v,\ y_v,\ v_x,\ v_y,\ \theta,\ B,\ \rho,\ \phi_1,\ \phi_2]^T \tag{5.15}$$

where B is the accelerometer bias. The dynamics of the state are:

$$\begin{cases} \dot{x}_v &= v_x \\ \dot{y}_v &= v_y \\ \dot{v}_x &= \cos\theta(A_x + B) - \sin\theta A_y \\ \dot{v}_y &= \sin\theta(A_x + B) + \cos\theta A_y \\ \dot{\theta} &= \omega \\ \dot{B} &= 0 \\ \dot{\rho} &= 0 \\ \dot{\phi}_1 &= 0 \\ \dot{\phi}_2 &= 0 \end{cases} \tag{5.16}$$

The analytic expression of the output is obtained starting from the expression $\beta = \pi + \phi_c - \theta_c$, where ϕ_c and θ_c are the bearing of the camera and its orientation in the global frame (see fig. 5.5). We compute the tangent of this expression, obtaining:

$$y \triangleq \tan\beta = \frac{y_c \cos\theta_c - x_c \sin\theta_c}{x_c \cos\theta_c + y_c \sin\theta_c} \tag{5.17}$$

where $(x_c,\ y_c)$ is the camera position in the global frame. We can express the previous output in terms of the state components by using the following equations: $x_c = x_v + \rho\cos(\theta + \phi_1)$, $y_c = y_v + \rho\sin(\theta + \phi_1)$ and $\theta_c = \theta + \phi_1 + \phi_2$.

5.5.1 System Canonization

We need to check if the system is in canonic form and, if not, we need to proceed as explained in appendix A.2. This is exactly what we have done in section 5.4.1. In particular, as in the case of calibrated sensors, we need to compute the space of functions $\mathcal{F}^2$ and we can check that the dimension of the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^2$ is 2. On the other hand, the functions h_1 and h_2 that we automatically select from $\mathcal{F}^2$, have now an analytic expression much more complex than in the case of calibrated sensors. Even if this fact does not prevent the application of the method in 4.2, for educational purposes, we prefer to proceed in a different manner. In particular, we apply the result stated in the remark 4 (in appendix A). Specifically, it is possible to check, by running algorithm 6, that the differential of the scalar $v_x^2 + v_y^2$ belongs to the observable codistribution. Hence, we can assume that the quantity $v_x^2 + v_y^2$ is a further system output. It is possible to check that the system defined by the state in (5.15), the dynamics in (5.16), the outputs $v_x^2 + v_y^2$ and the one in (5.17) is directly in canonic form.

5.5.2 Observability properties

We apply the method in 4.2. By comparing the dynamics in (5.16) with (2.1) we obtain:

$$g^0 = \begin{bmatrix} v_x \\ v_y \\ B\cos\theta \\ B\sin\theta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad f \triangleq f^1 = \begin{bmatrix} 0 \\ 0 \\ \cos\theta \\ \sin\theta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$g^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad g^2 = \begin{bmatrix} 0 \\ 0 \\ -\sin\theta \\ \cos\theta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

First step

In accordance with what we mentioned in 5.5.1, we can select h_1 and h_2 as follows:

$$h_1 = \frac{y_c\cos\theta_c - x_c\sin\theta_c}{x_c\cos\theta_c + y_c\sin\theta_c}$$
$$h_2 = v_x^2 + v_y^2$$

From (3.3) we obtain: $\mu_1^2 = \mu_2^1 = 0$, $\mu_1^1 =$

$$\frac{-\rho(x_v\cos(\phi_1+\theta) + y_v\sin(\phi_1+\theta)) - x_v^2 - y_v^2}{(\rho\cos(\phi_2) + x_v\cos(\phi_1+\phi_2+\theta) + y_v\sin(\phi_1+\phi_2+\theta))^2}$$

and $\mu_2^2 = 2v_y\cos\theta - 2v_x\sin\theta$, which is non singular (and diagonal).

Second step

We compute the inverse of the previous tensor. Because of the diagonal structure, we easily obtain: $\nu_1^2 = \nu_2^1 = 0$, $\nu_1^1 = \frac{1}{\mu_1^1}$ and $\nu_2^2 = \frac{1}{\mu_2^2}$. Additionally, we obtain from (3.6):

$$\mu_1^0 = \frac{\rho(v_y\cos(\phi_1+\theta) - v_x\sin(\phi_1+\theta)) + v_y x_v - v_x y_v}{(\rho\cos\phi_2 + x_v\cos(\phi_1+\phi_2+\theta) + y_v\sin(\phi_1+\phi_2+\theta))^2}$$
$$\mu_2^0 = 2B(v_x\cos\theta + v_y\sin\theta)$$

Finally, from (3.9) we obtain:

$$\hat{g}^0 = \begin{bmatrix} v_x \\ v_y \\ \frac{Bv_y}{v_y \cos\theta - v_x \sin\theta} \\ \frac{-Bv_x}{v_y \cos\theta - v_x \sin\theta} \\ \frac{\rho(v_y \cos(\phi_1+\theta) - v_x \sin(\phi_1+\theta)) + v_y x_v - v_x y_v}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The vector field $\hat{g}^1$ has all the entries null with the exception of the fifth entry that is

$$\frac{-(\rho\cos\phi_2 + x_v\cos(\phi_1+\phi_2+\theta) + y_v\sin(\phi_1+\phi_2+\theta))^2}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v}$$

The vector field $\hat{g}^2$ has all the entries null with the exception of the third and fourth ones, which are $\frac{-\sin\theta}{2v_y\cos\theta - 2v_x\sin\theta}$ and $\frac{\cos\theta}{2v_y\cos\theta - 2v_x\sin\theta}$, respectively.

Third step

We compute the differential of h_1 and h_2 (obtained at the first step) with respect to the state in (5.15). By a direct computation we obtain that the dimension of Ω_0 is 2. Additionally,we compute Ω_1. Its dimension is 3.

Fourth step

We have $\mathcal{L}_{\phi_0}\mathcal{L}_{g^j}h_i = \mathcal{L}_f\mu_i^j = 0$, $\forall i, j$. In order to check if we are in the special case considered by lemma 14 we need to compute ${}^0\phi_1, {}^1\phi_1$ and ${}^2\phi_1$, by using algorithm 4. We obtain that only the first two differ from the null vector and they are:

$${}^0\phi_1 = \begin{bmatrix} -\cos\theta \\ -\sin\theta \\ -\frac{\sin\theta(\rho(v_y\cos(\phi_1+\theta) - v_x\sin(\phi_1+\theta)) + v_y x_v - v_x y_v)}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v} \\ \frac{\cos\theta(\rho(v_y\cos(\phi_1+\theta) - v_x\sin(\phi_1+\theta)) + v_y x_v - v_x y_v)}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$${}^1\phi_1 = \begin{bmatrix} 0 \\ 0 \\ \frac{\sin\theta(\rho\cos\phi_2 + x\cos(\phi_1+\phi_2+\theta) + y_v\sin(\phi_1+\phi_2+\theta))^2}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v} \\ -\frac{\cos\theta(\rho\cos\phi_2 + x_v\cos(\phi_1+\phi_2+\theta) + y_v\sin(\phi_1+\phi_2+\theta))^2}{x_v^2 + \rho\cos(\phi_1+\theta)x_v + y_v^2 + \rho\sin(\phi_1+\theta)y_v} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We have $\mathcal{L}_{^1\phi_1}\mathcal{L}_{g^2}h_2 = \mathcal{L}_{^1\phi_1}\mu_2^2 = -2\frac{(r\cos\phi_2+x_v\cos(\phi_1+\phi_2+\theta)+y_v\sin(\phi_1+\phi_2+\theta))^2}{x_v^2+r\cos(\phi_1+\theta)x_v+y_v^2+r\sin(\phi_1+\theta)y_v}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 14 and we need to continue with step 5.

Fifth step

We compute the three-index tensor in (3.11). We do not provide here the expression of its components, for brevity sake. We only mention that the non-vanishing components are the following eleven: $\mathcal{T}_1^{00}$, $\mathcal{T}_2^{00}$, $\mathcal{T}_1^{01}$, $\mathcal{T}_2^{02}$, $\mathcal{T}_1^{10}$, $\mathcal{T}_2^{10}$, $\mathcal{T}_1^{20}$, $\mathcal{T}_2^{20}$, $\mathcal{T}_1^{11}$, $\mathcal{T}_2^{12}$ and $\mathcal{T}_2^{22}$.

Sixth step

By using algorithm 5 and the vector fields ${}^0\phi_1,{}^1\phi_1$ and ${}^2\phi_1$ previously computed, we can compute Ω_2. Its dimension is 7. We obtain that the differentials of all the components $\mathcal{T}$ belong to Ω_2 with the exception of $\mathcal{T}_1^{00}$, $\mathcal{T}_1^{01}$, $\mathcal{T}_1^{10}$, $\mathcal{T}_1^{20}$ and $\mathcal{T}_1^{11}$. Hence, we need to compute Ω_3 by using algorithms 4 and 5. We do not provide all the steps. It is possible to check that also the differentials of the remaining components of $\mathcal{T}$ belong to Ω_3, whose dimension is 8. Therefore, we have $m' = 3$.

Seventh step

By an explicit computation (by using the subsequent step of both algorithms 4 and 5), we obtain $\Omega_4 = \Omega_3$. Hence, algorithm 5 has coverged and $\Omega^* = \Omega_3$.

Eighth step

By computing the distribution orthogonal to the codistribution Ω_3 we can find the continuous symmetry that characterizes the unobservable space [37]. By an explicit computation we obtain the vector: $[y_v,\ -x_v,\ v_y,\ -v_x,\ 1,\ 0,\ 0,\ 0,\ 0]^T$, which corresponds to a rotation around the vertical axis. Note that, the absolute scale is invariant with respect to a rotation around the vertical axis meaning that it can be estimated. This also holds for the sensor suit speed in the local frame.

We remark that we obtain the same result that holds in the case of calibrated sensors. This means that the sensors, even if not calibrated, provide enough information to perform their self-calibration. Additionally, as in the case of calibrated sensors, we remark that, the invariance that corresponds to the continuous symmetry detected by our observability analysis, would be present also by having a complete IMU, i.e., when the IMU provides the acceleration along the two axes and the angular speed. This means that the information obtained by fusing the measurements from a camera with the measurements from a complete IMU is redundant.

5.6 Visual-inertial sensor fusion: the 3D case with calibrated sensors

We consider again a sensor suit that consists of inertial sensors (IMU) and a bearing sensor (camera). However, we now investigate the $3D$ case. Without loss of generality, we define the sensor suit local frame as the IMU frame. The inertial sensors measure the sensor suit acceleration and the angular speed. In $3D$, the acceleration and the angular speed are three dimensional vectors. The camera provides the bearing angles of the features in its own local frame.

We derive the observability properties in the case when the IMU only provides the acceleration along a single axis (instead of three). In other words, we will consider the other components of the acceleration and the angular speed as unknown inputs. Therefore, the system is characterized by five unknown inputs. Finally, we consider the extreme case of a single point feature.

In general, the camera frame does not coincide with the IMU frame. Additionally, the measurements provided by the IMU are in general biased. We start our investigation in the simple case when the camera frame coincides with the IMU frame and the inertial measurements are unbiased. Then, in section 5.7, we relax both these assumptions by also assuming that the camera is extrinsically uncalibrated (i.e., the transformation between the IMU frame and the camera frame is unknown).

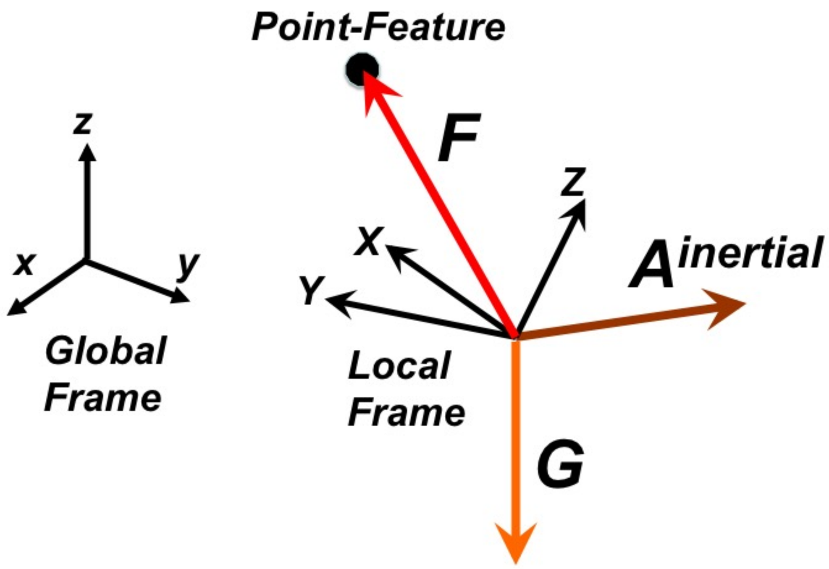

Figure 5.6: Visual-inertial sensor fusion in $3D$ with calibrated sensors and in the case of a single point feature.

Our system can be characterized by the following state (see fig 5.6 for an illustration):

$$X \triangleq [F_x,\ F_y,\ F_z,\ V_x,\ V_y,\ V_z,\ q_t,\ q_x,\ q_y,\ q_z,\ g]^T \tag{5.18}$$

where $F = [F_x,\ F_y,\ F_z]$ is the position of the point feature in the local frame, $V = [V_x,\ V_y,\ V_z]$ is the sensor suit speed in the local frame, $q = q_t + q_x i + q_y j + q_z k$ is the unit quaternion that describes the transformation change between the global and the local frame and g is the

magnitude of the gravity (that is assumed unknown). The dynamics are:

$$\begin{bmatrix} \dot{F} & = -\Omega \wedge F - V \\ \dot{V} & = -\Omega \wedge V + A + G \\ \dot{q} & = \frac{1}{2} q \Omega_q \\ \dot{g} & = 0 \end{bmatrix} \tag{5.19}$$

where $\Omega \triangleq [\Omega_x \ \Omega_y \ \Omega_z]$ is the unknown angular speed of the camera, Ω_q is the imaginary quaternion associated with Ω, i.e., $\Omega_q \triangleq \Omega_x i + \Omega_y j + \Omega_z k$, G is the gravity in the local frame and A is the acceleration that would be perceived by a noiseless tri-axis accelerometer. In other words, since the accelerometer also perceives the gravity, $A^{inertial} \triangleq A + G$ is the inertial acceleration expressed in the local frame. Without loss of generality, we assume that the accelerometer provides the third component of A. In other words, the first two components of A act as unknown inputs.

The monocular camera provides the position of the feature in the local frame (F) up to a scale. Hence, it provides the ratios of the components of F:

$$h_{cam}(X) \triangleq [h_u, \ h_v]^T = \left[\frac{F_x}{F_z}, \ \frac{F_y}{F_z}\right]^T \tag{5.20}$$

We have also to consider the constraint $q^* q = 1$. This provides the further observation:

$$h_{const}(X) \triangleq h_q = q^* q \tag{5.21}$$

We consider the system characterized by the state in (5.18), the dynamics in (5.19) and the three outputs h_u, h_v and h_q in (5.20) and (5.21) under the assumptions that the IMU only provides the acceleration (inertial and gravitational) along the $z-$axis of the local frame. Hence, we have a single known input ($m_u = 1$) that is $u = u_1 = A_z$. Additionally, we have five unknown inputs ($m_w = 5$) that are: $w_1 = \Omega_x$, $w_2 = \Omega_y$, $w_3 = \Omega_z$, $w_4 = A_x$ and $w_5 = A_y$. By comparing (5.19) with (2.1) we obtain:

$$g^0 = \begin{bmatrix} -V_x \\ -V_y \\ -V_z \\ -2g(q_t q_y - q_x q_z) \\ 2g(q_t q_x + q_y q_z) \\ g(q_t^2 - q_x^2 - q_y^2 + q_z^2) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad f \triangleq f^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$g^1 = \begin{bmatrix} 0 \\ F_z \\ -F_y \\ 0 \\ V_z \\ -V_y \\ -q_x/2 \\ q_t/2 \\ q_z/2 \\ -q_y/2 \\ 0 \end{bmatrix}, \quad g^2 = \begin{bmatrix} -F_z \\ 0 \\ F_x \\ -V_z \\ 0 \\ V_x \\ -q_y/2 \\ -q_z/2 \\ q_t/2 \\ q_x/2 \\ 0 \end{bmatrix}, \quad g^3 = \begin{bmatrix} F_y \\ -F_x \\ 0 \\ V_y \\ -V_x \\ 0 \\ -q_z/2 \\ q_y/2 \\ -q_x/2 \\ q_t/2 \\ 0 \end{bmatrix}$$

$$g^4 = [0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T$$
$$g^5 = [0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T$$

5.6.1 System Canonization

We need to check if the system is in canonic form and, if not, we need to proceed as explained in appendix A.2. This is exactly what we have done in sections 5.4.1 and 5.5.1. In particular, we need to compute the space of functions $\mathcal{F}^3$ and we can check that the dimension of the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^3$ is 5. On the other hand, the functions $h_1,\ h_2,\ h_3,\ h_4$ and h_5 that we automatically select from $\mathcal{F}^3$, have a complex analytic expression. Even if this fact does not prevent the application of the method in 4.2, for educational purposes, we prefer to proceed in a different manner. In particular, we apply the result stated in remark 3 (in appendix A). Specifically, it is possible to check that the functions $h_1 = F_x,\ h_2 = F_y,\ h_3 = V_x,\ h_4 = V_y$ and $h_5 = V_z$ belong to $\mathcal{F}^3$ (the first two functions also belong to $\mathcal{F}^2$). Additionally, it is possible to check that the system defined by the state in (5.18), the dynamics in (5.19) and the outputs $h_1,\ h_2,\ h_3,\ h_4$ and h_5 is directly in canonic form. Note that we applied the result stated by the remark 3 (in appendix A), with $k = 0$, because the five functions only depend on the original state.

5.6.2 Observability properties

We apply the method in 4.2 to obtain the observability properties of the system characterized by the state in (5.18), the dynamics in (5.19) and the outputs $h_1 = F_x,\ h_2 = F_y,\ h_3 = V_x,\ h_4 = V_y,$ $h_5 = V_z$ and $h_q = q_t^2 + q_x^2 + q_y^2 + q_z^2$.

First step

In accordance with what we mentioned in 5.6.1, the system is in canonic form. From (3.3) we obtain:

$$\mu = \begin{bmatrix} 0 & -F_z & F_y & 0 & 0 \\ F_z & 0 & -F_x & 0 & 0 \\ 0 & -V_z & V_y & 1 & 0 \\ V_z & 0 & -V_x & 0 & 1 \\ -V_y & V_x & 0 & 0 & 0 \end{bmatrix}$$

where the upper index corresponds to the column and the lower index to the line. This tensor is non-singular.

Second step

We compute the inverse of the previous tensor. We easily obtain

$$\nu = \frac{1}{F_z(F_xV_y - F_yV_x)} \begin{bmatrix} -F_xV_x & -F_yV_x & 0 & 0 & -F_xF_z \\ -F_xV_y & -F_yV_y & 0 & 0 & -F_yF_z \\ -F_zV_x & -F_zV_y & 0 & 0 & -F_z^2 \\ -V_y(F_xV_z - F_zV_x) & F_zV_y^2 - F_yV_yV_z & F_z(F_xV_y - F_yV_x) & 0 & -F_z(F_yV_z - F_zV_y) \\ F_xV_xV_z - F_zV_x^2 & V_x(F_yV_z - F_zV_y) & 0 & F_z(F_xV_y - F_yV_x) & F_z(F_xV_z - F_zV_x) \end{bmatrix}$$

Additionally, we obtain from (3.6):

$$\begin{aligned} \mu_1^0 &= -\,V_x \\ \mu_2^0 &= -\,V_y \\ \mu_3^0 &= -\,2g(q_tq_y - q_xq_z) \\ \mu_4^0 &= 2g(q_tq_x + q_yq_z) \\ \mu_5^0 &= g(q_t^2 - q_x^2 - q_y^2 + q_z^2) \end{aligned}$$

Finally, from (3.9) we obtain:

$$\hat{g}^1 = \frac{1}{2\eta} \begin{bmatrix} 2F_z(F_xV_y - F_yV_x) \\ 0 \\ 2F_xF_yV_x - 2F_x^2V_y \\ 0 \\ 0 \\ 0 \\ F_xV_xq_x + F_xV_yq_y + F_zV_xq_z \\ F_xV_yq_z - F_xV_xq_t - F_zV_xq_y \\ F_zV_xq_x - F_xV_xq_z - F_xV_yq_t \\ F_xV_xq_y - F_zV_xq_t - F_xV_yq_x \\ 0 \end{bmatrix}, \quad \hat{g}^3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\hat{g}^2 = \frac{1}{2\eta} \begin{bmatrix} 0 \\ 2F_z(F_xV_y - F_yV_x) \\ 2F_y^2V_x - 2F_xF_yV_y \\ 0 \\ 0 \\ 0 \\ F_yV_xq_x + F_yV_yq_y + F_zV_yq_z \\ F_yV_yq_z - F_yV_xq_t - F_zV_yq_y \\ F_zV_yq_x - F_yV_xq_z - F_yV_yq_t \\ F_yV_xq_y - F_zV_yq_t - F_yV_yq_x \\ 0 \end{bmatrix}, \quad \hat{g}^4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \hat{g}^5 = \frac{F_z}{2\eta} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2F_xV_y - 2F_yV_x \\ F_xq_x + F_yq_y + F_zq_z \\ F_yq_z - F_xq_t - F_zq_y \\ F_zq_x - F_xq_z - F_yq_t \\ F_xq_y - F_zq_t - F_yq_x \\ 0 \end{bmatrix}$$

where $\eta = F_z(F_xV_y - F_yV_x)$. We do not provide the expression of g^0 for the brevity sake.

Third step

We compute the differential of h_1, h_2, h_3, h_4, h_5 and h_q with respect to the state in (5.18). By a direct computation we obtain that the dimension of Ω_0 is 6. Additionally, we obtain $\Omega_0 = \Omega_1$.

Fourth step

We have $\mathcal{L}_{\phi_0}\mathcal{L}_{g^1}h_4 = \mathcal{L}_f \mu_4^1 = 1 \neq 0$. This suffices to conclude that the considered system is not in the special case considered by lemma 14 and we need to continue with step 5.

Fifth step

We compute the three-index tensor in (3.11). We remind the reader that we can consider the lower index as a Latin index since the components of the tensor when this index is zero vanish. Since the Latin index takes the values $1, \cdots, 5$ and the Greek indexes $0, \cdots, 5$, this tensor has 180 components that can be different from zero. To display these components, we provide separately $\mathcal{T}_i^{1j}$, $\mathcal{T}_i^{2j}$, $\mathcal{T}_i^{3j}$, $\mathcal{T}_i^{4j}$, $\mathcal{T}_i^{5j}$. We do not provide $\mathcal{T}_i^{00}$, $\mathcal{T}_i^{0j}$, $\mathcal{T}_i^{j0}$ for the brevity sake. We have:

$$
\mathcal{T}_i^{1j} = \frac{1}{\eta}\begin{bmatrix} -(F_x^2 V_y)/F_z & -(F_x F_y V_y)/F_z & 0 & 0 & -F_x F_y \\ (V_x(F_x^2+F_z^2))/F_z & (V_y F_z^2 + F_x F_y V_x)/F_z & 0 & 0 & F_x^2+F_z^2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad
\mathcal{T}_i^{2j} = \frac{1}{\eta}\begin{bmatrix} -(V_x F_z^2 + F_x F_y V_y)/F_z & -(V_y(F_y^2+F_z^2))/F_z & 0 & 0 & -F_y^2-F_z^2 \\ (F_x F_y V_x)/F_z & (F_y^2 V_x)/F_z & 0 & 0 & F_x F_y \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\mathcal{T}_i^{3j} = \frac{1}{\eta}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ F_z V_x & F_z V_y & 0 & 0 & F_z^2 \\ -F_x V_y & -F_y V_y & 0 & 0 & -F_y F_z \end{bmatrix}
$$

$$
\mathcal{T}_i^{4j} = \frac{1}{\eta}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -F_z V_x & -F_z V_y & 0 & 0 & -F_z^2 \\ 0 & 0 & 0 & 0 & 0 \\ F_x V_x & F_y V_x & 0 & 0 & F_x F_z \end{bmatrix}
$$

$$
\mathcal{T}_i^{5j} = \frac{1}{\eta}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ F_x V_y & F_y V_y & 0 & 0 & F_y F_z \\ -F_x V_x & -F_y V_x & 0 & 0 & -F_x F_z \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

Sixth step

We need to compute Ω_2 and, in order to do this, we need to compute ${}^0\phi_1, {}^1\phi_1, {}^2\phi_1, {}^3\phi_1, {}^4\phi_1$ and ${}^5\phi_1$, by using algorithm 4. We obtain that ${}^3\phi_1$ and ${}^4\phi_1$ are null. The remaining four are: ${}^0\phi_1 = \frac{1}{\eta}$

$$\begin{bmatrix} 0 \\ 0 \\ F_z(F_xV_y - F_yV_x) \\ g(F_yq_t^2 - F_yq_x^2 - F_yq_y^2 + F_yq_z^2)F_z - (F_yV_y^2 + F_xV_xV_y) \\ (F_xV_x^2 + F_yV_yV_x) - F_zg(F_xq_t^2 - F_xq_x^2 - F_xq_y^2 + F_xq_z^2) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$${}^1\phi_1 = \frac{1}{\eta}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -F_xV_y \\ F_xV_x \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, {}^2\phi_1 = \frac{1}{\eta}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -F_yV_y \\ F_yV_x \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, {}^5\phi_1 = \frac{1}{\eta}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -F_yF_z \\ F_xF_z \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence, by using algorithm 5 we obtain Ω_2. Its dimension is 8. We obtain that the differentials of all the components of $\mathcal{T}$ belong to Ω_2 with the exception of $\mathcal{T}_3^{00}$, $\mathcal{T}_4^{00}$, $\mathcal{T}_5^{00}$, $\mathcal{T}_3^{10}$, $\mathcal{T}_4^{10}$, $\mathcal{T}_5^{10}$, $\mathcal{T}_3^{20}$, $\mathcal{T}_4^{20}$, $\mathcal{T}_5^{20}$, $\mathcal{T}_3^{50}$, $\mathcal{T}_4^{50}$ and $\mathcal{T}_5^{50}$.

We need to compute Ω_3 by using algorithms 4 and 5. We do not provide all the steps. We obtain that the dimension of Ω_3 is 10. Additionally, we obtain that the differentials of all the components of $\mathcal{T}$ belong to Ω_3. Hence, we have $m' = 2$.

Seventh step

We compute Ω_4 by using algorithms 4 and 5. We obtain $\Omega_4 = \Omega_3$. Hence, algorithm 5 has converged to the codistribution $\Omega^* = \Omega_3$.

Eighth step

By computing the distribution orthogonal to the codistribution Ω^* we can find the continuous symmetry that characterizes the unobservable space [37]. By an explicit computation we obtain the vector: $[0,\ 0,\ 0,\ 0,\ 0,\ 0,\ -q_z,\ -q_y,\ q_x,\ q_t,\ 0]^T$, which corresponds to a rotation around the vertical axis (the axis aligned with the gravity).

We remark that the invariance that corresponds to the continuous symmetry detected by our observability analysis, would be present also by having a complete IMU, i.e., when the IMU provides the acceleration along the three axes and the three components of the angular speed.

This means that the information obtained by fusing the measurements from a camera with the measurements from a complete IMU is redundant.

5.7 Visual-inertial sensor fusion: the 3D case with uncalibrated sensors

We now consider the same sensor fusion problem considered in the previous section, but we assume that the camera is not extrinsically calibrated and the inertial measurements are biased. We need to include in the state the inertial bias and the parameters that describe the transformation change between the camera and the IMU. We remind the reader that the IMU only consists of a single axis accelerometer.

We assume that the local frame coincides with the frame attached to the single-axis accelerometer and, without loss of generality, we assume that this local frame has its z-axis coincident with the axis of the accelerometer. The position of the camera optical center in the local frame will be denoted by $P_c = [X_c,\ Y_c,\ Z_c]^T$ and the camera orientation will be characterized through the three Euler angles α, β, γ. Specifically, a vector with orientation $\hat{\tau}$ in the local frame, will have the orientation $R\hat{\tau}$ in the camera frame, where $R = R(\alpha, \beta, \gamma) = R^z_\alpha R^x_\beta R^z_\gamma$ and R^z_η and R^x_η rotates the unit vector $\hat{\tau}$ clockwise through the angle η about the $z-$axis and the $x-$axis, respectively. The vector P_c and the three angles α, β, γ, characterize the extrinsic camera calibration and are assumed to be unknown. As in the previous section, we consider the extreme case when a single point feature is available and we denote its position in the camera frame with ${}^cF \triangleq [{}^cF_x\ {}^cF_y\ {}^cF_z]^T$. We characterize our system through the following state:

$$X \triangleq \left[{}^cF^T,\ V^T,\ q_t,\ q_x,\ q_y,\ q_z,\ g, P_c^T,\ \alpha,\ \beta,\ \gamma,\ B\right]^T \tag{5.22}$$

where B is the accelerometer bias. Note that the position of the point feature is expressed in the camera frame, while the speed, $V = [V_x,\ V_y,\ V_z]^T$, is in the local frame. Additionally, $q = q_t + q_x i + q_y j + q_z k$ is the unit quaternion that describes the transformation change between the global and the local frame. The dynamics are:

$$\begin{cases} {}^c\dot{F} = -{}^c\Omega \wedge^c F - R(V + \Omega \wedge P_c) \\ \dot{V} = -\Omega \wedge V + A + G \\ \dot{q} \ = \frac{1}{2} q \Omega_q \\ \dot{P}_c = [0, 0, 0]^T \quad \dot{g} = \dot{B} = \dot{\alpha} = \dot{\beta} = \dot{\gamma} = 0 \end{cases} \tag{5.23}$$

where ${}^c\Omega$ is the angular speed expressed in the camera frame, which is related to the one expressed in the local frame through the rotation matrix R, ${}^c\Omega = R\Omega$. As in the previous section, the accelerometer provides the third component of the acceleration perceived in the local frame, which includes the gravity. We assume that the measurements are affected by the bias B.

The monocular camera provides the position of the feature in the camera frame (${}^cF$) up to a scale. Hence, it provides the ratios of the components of ${}^cF$:

$$h_{cam}(X) \triangleq [h_u,\ h_v]^T = \left[\frac{{}^cF_x}{{}^cF_z},\ \frac{{}^cF_y}{{}^cF_z}\right]^T \tag{5.24}$$

Additionally, as in the previous section, we have also to consider the constraint $q^*q = 1$. This provides the further observation given in (5.21).

We consider the system characterized by the state in (5.22), the dynamics in (5.23) and the three outputs h_u, h_v and h_q in (5.24) and (5.21) under the assumptions that the IMU only provides the acceleration (inertial and gravitational) along the $z-$axis of the local frame. Hence, we have a single known input ($m_u = 1$) that is $u = u_1 = A_z$. Additionally, we have five unknown inputs ($m_w = 5$) that are: $w_1 = \Omega_x$, $w_2 = \Omega_y$, $w_3 = \Omega_z$, $w_4 = A_x$ and $w_5 = A_y$.

By comparing (5.23) with (2.1) we obtain the expressions of the vector fields: g^0, g^1, g^2, g^3, g^4, g^5 and f. In particular:

- g^0 is obtained by setting to zero all the components of both A and Ω in (5.23), with the exception of the third component of A, which is set equal to $-B$;

- f is obtained by removing g^0 from (5.23) and then by setting $A_x = A_y = \Omega_x = \Omega_y = \Omega_z = 0$ and $A_z = 1$;

- g^j (for $j = 1, \cdots, 5$) are obtained by removing g^0 from (5.23) and then by setting to zero all the components of both A and Ω with the exception of one of them depending on j (e.g., to obtain g^1 we set $A_x = A_y = A_z = \Omega_y = \Omega_z = 0$ and $\Omega_x = 1$).

5.7.1 System Canonization

We need to check if the system is in canonic form and, if not, we need to proceed as explained in appendix A. This is exactly what we have done in section 5.6.1. By proceeding with the procedure described in appendix A it is possible to set the system in canonic form. On the other hand, the analytic expressions of the quantities computed by the application of this procedure, are complex. For educational purposes, we prefer to proceed differently. In particular, by using algorithm 6, it is possible to prove that the differential of the following scalar functions, that only depend on the components of the state in 5.22, belong to the observable codistribution: ${}^cF_x$, ${}^cF_y$, ${}^cF_z$, $V_x^2 + V_y^2$, V_z, $\frac{V_y \cos\gamma - V_x \sin\gamma}{V_x \cos\gamma + V_y \sin\gamma}$, α, β and B. Hence, we can use these functions as system outputs. In particular, we find that the system is directly in canonic form by using the following outputs: $h_1 = {}^cF_x$, $h_2 = {}^cF_y$, $h_3 = \frac{V_y \cos\gamma - V_x \sin\gamma}{V_x \cos\gamma + V_y \sin\gamma}$, $h_4 = V_x^2 + V_y^2$ and $h_5 = V_z$. Therefore, we use these five scalar functions for the m_w scalar functions that we have to select, in accordance with the first step of the method in section 4.2. Before proceeding, in order to further simplify the analytic computation, we can remove α, β and B from the state. In particular, we can set their values to zero. Hence we refer to the following state:

$$X \triangleq \left[{}^cF^T,\ V^T,\ q_t,\ q_x,\ q_y,\ q_z,\ g, P_c^T,\ \gamma\right]^T \tag{5.25}$$

whose dynamics are easily obtained from (5.23). Note that now the matrix R only describes a rotation of γ about the accelerometer axis. By comparing these dynamics with (2.1) we obtain:

$$g^0 = \begin{bmatrix} -V_x \cos\gamma - V_y \sin\gamma \\ V_x \sin\gamma - V_y \cos\gamma \\ -V_z \\ -2g(q_t q_y - q_x q_z) \\ 2g(q_t q_x + q_y q_z) \\ g(q_t^2 - q_x^2 - q_y^2 + q_z^2) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad f \triangleq f^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$g^1 = \begin{bmatrix} \sin\gamma({}^cF_z + Z_c) \\ \cos\gamma({}^cF_z + Z_c) \\ -Y_c -{}^c F_y \cos\gamma -{}^c F_x \sin\gamma \\ 0 \\ V_z \\ -V_y \\ -q_x/2 \\ q_t/2 \\ q_z/2 \\ -q_y/2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$g^2 = \begin{bmatrix} -\cos\gamma({}^cF_z + Z_c) \\ \sin\gamma({}^cF_z + Z_c) \\ X_c +{}^c F_x \cos\gamma -{}^c F_y \sin\gamma \\ -V_z \\ 0 \\ V_x \\ -q_y/2 \\ -q_z/2 \\ q_t/2 \\ q_x/2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad g^3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$g^4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad g^5 = \begin{bmatrix} {}^cF_y + Y_c \cos\gamma - X_c \sin\gamma \\ -{}^cF_x - X_c \cos\gamma - Y_c \sin\gamma \\ 0 \\ V_y \\ -V_x \\ 0 \\ -q_z/2 \\ q_y/2 \\ -q_x/2 \\ q_t/2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

5.7.2 Observability properties

We apply the method in 4.2 to obtain the observability properties of the system characterized by the state in (5.25), the dynamics in (5.23) (without α, β and B) and the outputs:

$$\begin{aligned} h_1 &= {}^cF_x \\ h_2 &= {}^cF_y \\ h_3 &= \frac{V_y \cos\gamma - V_x \sin\gamma}{V_x \cos\gamma + V_y \sin\gamma} \\ h_4 &= V_x^2 + V_y^2 \\ h_5 &= V_z \end{aligned}$$

and $h_z =^c F_z$ and $h_q = q_t^2 + q_x^2 + q_y^2 + q_z^2$.

First step

In accordance with what we mentioned in 5.7.1, the system is directly in canonic form. From (3.3) we obtain: $\mu_i^j =$

$$\begin{bmatrix} \sin\gamma({}^cF_z + Z_c) & -\cos\gamma({}^cF_z + Z_c) & 0 & 0 & {}^cF_y + Y_c \cos\gamma - X_c \sin\gamma \\ \cos\gamma({}^cF_z + Z_c) & \sin\gamma({}^cF_z + Z_c) & 0 & 0 & {}^cF_y + Y_c \cos\gamma - X_c \sin\gamma \\ \frac{V_x V_z}{(V_x \cos\gamma + V_y \sin\gamma)^2} & \frac{V_y V_z}{(V_x \cos\gamma + V_y \sin\gamma)^2} & \frac{-V_y}{(V_x \cos\gamma + V_y \sin\gamma)^2} & \frac{V_x}{(V_x \cos\gamma + V_y \sin\gamma)^2} & \frac{-(V_x^2 + V_y^2)}{(V_x \cos\gamma + V_y \sin\gamma)^2} \\ 2V_y V_z & -2V_x V_z & 2V_x & 2V_y & 0 \\ -V_y & V_x & 0 & 0 & 0 \end{bmatrix}$$

where the upper index corresponds to the column and the lower index to the line. This tensor is non-singular.

Second step

We compute the inverse of the previous tensor. We do not provide the expression for the brevity sake. Additionally, we obtain from (3.6):

$$
\begin{aligned}
\mu_1^0 =& - V_x \cos\gamma - V_y \sin\gamma \\
\mu_2^0 =& V_x \sin\gamma - V_y \cos\gamma \\
h_3^0 =& \frac{V_x(2gq_tq_x + 2gq_yq_z) + V_y(2gq_tq_y - 2gq_xq_z)}{(V_x\cos\gamma + V_y\sin\gamma)^2} \\
\mu_4^0 =& 4V_yg(q_tq_x + q_yq_z) - 4V_xg(q_tq_y - q_xq_z) \\
\mu_5^0 =& g(q_t^2 - q_x^2 - q_y^2 + q_z^2)
\end{aligned}
$$

Finally, from (3.9) we obtain the expressions of $\hat{g}^0$, $\hat{g}^1$, $\hat{g}^2$, $\hat{g}^3$, $\hat{g}^4$ and $\hat{g}^5$. For the brevity sake, we only provide the expressions of $\hat{g}^3$ and $\hat{g}^4$, which are simple.

$$
\hat{g}^3 = \frac{(V_x\cos\gamma + V_y\sin\gamma)^2}{V_x^2 + V_y^2} \times
$$

$$
[0,\ 0,\ 0,\ -V_y,\ V_x,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T
$$

$$
\hat{g}^4 = \frac{1}{2(V_x^2 + V_y^2)} \times
$$

$$
[0,\ 0,\ 0,\ V_x,\ V_y,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T
$$

Third step

We compute the differential of h_1, h_2, h_3, h_4, h_5, h_z and h_q with respect to the state in (5.25). By a direct computation we obtain that the dimension of Ω_0 is 7. Additionally, we compute Ω_1 and we obtain $\Omega_1 = \Omega_0$.

Fourth step

We have $\mathcal{L}_{\phi_0}\mathcal{L}_{g^1}h_4 = \mathcal{L}_f\mu_4^1 = 2V_y \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 14 and we need to continue with step 5.

Fifth step

We compute the three-index tensor in (3.11). We remind the reader that we can consider the lower index as a Latin index since the components of the tensor when this index is zero vanish. Since the Latin index takes the values $1, \cdots, 5$ and the Greek indexes $0, \cdots, 5$, this tensor has 180 components that can differ from zero. For the brevity sake, we do not provide here the analytic expressions of its components. We only mention which are the non-vanishing components. They are the following 63: $\mathcal{T}_1^{00}$, $\mathcal{T}_2^{00}$, $\mathcal{T}_3^{00}$, $\mathcal{T}_4^{00}$, $\mathcal{T}_5^{00}$, $\mathcal{T}_1^{10}$, $\mathcal{T}_2^{10}$, $\mathcal{T}_3^{10}$, $\mathcal{T}_4^{10}$, $\mathcal{T}_5^{10}$, $\mathcal{T}_1^{20}$, $\mathcal{T}_2^{20}$, $\mathcal{T}_3^{20}$, $\mathcal{T}_4^{20}$, $\mathcal{T}_5^{20}$, $\mathcal{T}_1^{30}$, $\mathcal{T}_2^{30}$, $\mathcal{T}_4^{30}$, $\mathcal{T}_5^{30}$, $\mathcal{T}_1^{40}$, $\mathcal{T}_2^{40}$, $\mathcal{T}_3^{40}$, $\mathcal{T}_5^{40}$, $\mathcal{T}_3^{50}$, $\mathcal{T}_4^{50}$, $\mathcal{T}_5^{50}$, $\mathcal{T}_1^{01}$, $\mathcal{T}_1^{02}$, $\mathcal{T}_1^{05}$, $\mathcal{T}_2^{01}$, $\mathcal{T}_2^{02}$, $\mathcal{T}_2^{05}$, $\mathcal{T}_1^{11}$, $\mathcal{T}_1^{12}$, $\mathcal{T}_1^{15}$, $\mathcal{T}_2^{11}$, $\mathcal{T}_2^{12}$, $\mathcal{T}_2^{15}$, $\mathcal{T}_1^{21}$, $\mathcal{T}_1^{22}$, $\mathcal{T}_1^{25}$, $\mathcal{T}_2^{21}$, $\mathcal{T}_2^{22}$, $\mathcal{T}_2^{25}$, $\mathcal{T}_3^{33}$, $\mathcal{T}_3^{34}$, $\mathcal{T}_4^{33}$, $\mathcal{T}_4^{31}$, $\mathcal{T}_4^{32}$, $\mathcal{T}_4^{35}$, $\mathcal{T}_5^{31}$, $\mathcal{T}_5^{32}$, $\mathcal{T}_5^{35}$, $\mathcal{T}_3^{41}$, $\mathcal{T}_3^{42}$, $\mathcal{T}_3^{43}$, $\mathcal{T}_3^{45}$, $\mathcal{T}_4^{44}$, $\mathcal{T}_5^{45}$, $\mathcal{T}_3^{51}$, $\mathcal{T}_3^{52}$, $\mathcal{T}_3^{55}$ and $\mathcal{T}_4^{55}$.

Sixth step

We need to compute Ω_2 and, in order to do this, we need to compute ${}^0\phi_1, {}^1\phi_1, {}^2\phi_1, {}^3\phi_1, {}^4\phi_1$ and ${}^5\phi_1$, by using algorithm 4. We obtain that ${}^3\phi_1$ and ${}^4\phi_1$ are null. We do not provide the expression of the remaining four. By using algorithm 5 we obtain Ω_2. Its dimension is 10. We obtain that the differentials of the following 19 components of $\mathcal{T}$ belong to Ω_2: $\mathcal{T}_1^{03}$, $\mathcal{T}_2^{03}$, $\mathcal{T}_5^{03}$,

$\mathcal{T}_1^{04}$, $\mathcal{T}_2^{04}$, $\mathcal{T}_5^{04}$, $\mathcal{T}_3^{33}$, $\mathcal{T}_3^{34}$, $\mathcal{T}_4^{33}$, $\mathcal{T}_3^{43}$, $\mathcal{T}_4^{44}$, $\mathcal{T}_5^{45}$, $\mathcal{T}_4^{55}$, $\mathcal{T}_5^{31}$, $\mathcal{T}_5^{32}$, $\mathcal{T}_5^{35}$, $\mathcal{T}_3^{51}$, $\mathcal{T}_3^{52}$ and $\mathcal{T}_3^{55}$. Hence, the differentials of $63 - 19 = 44$ components of $\mathcal{T}$ do not belong to Ω_2.
We compute Ω_3 by using the subsequent steps of algorithms 4 and 5. Its dimension is 13. Additionally, the differentials of all the components of $\mathcal{T}$ belong to Ω_3. Hence, $m' = 3$;

Seventh step

We need to compute Ω_4 by using algorithms 4 and 5. We do not provide all the steps. We obtain that $\Omega_4 = \Omega_3$. Hence, algorithm 5 has converged to the codistribution $\Omega^* = \Omega_3$.

Eighth step

By computing the distribution orthogonal to the codistribution Ω^* we can find the continuous symmetries that characterize the unobservable space [37]. By an explicit computation we obtain the following two vectors:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -q_z/2 \\ -q_y/2 \\ q_x/2 \\ q_t/2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ -V_y \\ V_x \\ 0 \\ q_z/2 \\ -q_y/2 \\ q_x/2 \\ -q_t/2 \\ 0 \\ -Y_c \\ X_c \\ 0 \\ 1 \end{bmatrix}$$

The former corresponds to a rotation around the vertical axis (the axis aligned with the gravity). This was the only continuous symmetry that characterized the case with calibrated sensors, discussed in the previous section. It is non surprising that this symmetry remains. The latter corresponds to a rotation around the accelerometer axis.

We conclude this section by remarking that, if the camera is not extrinsically calibrated, an internal symmetry arises. As a result, it is not possible to distinguish all the physical quantities rotated around the accelerometer axis, independently of the accomplished trajectory. This means that, in this setting, it is not possible to fully perceive self-motion. If an additional inertial sensor is introduced, the latter symmetry is broken, provided that this additional sensor is not aligned with the accelerometer.

We summarize the results of the last two sections as follows:

In the visual-inertial sensor fusion problem with only two inertial sensors, the observability properties are the same as in the standard case provided that the two inertial sensors are along two distinct axes and with at least one of them that is an accelerometer. In other words, a sensor suit constituted of a monocular camera and two singe-axis inertial sensors, is able to perceive self-motion as a sensor suit constitutes by a monocular camera and a complete inertial measurement unit. This holds even in the most challenging scenario, i.e., in the case

of unknown camera-inertial sensor transformation, unknown magnitude of the gravity, unknown biases and single point feature available. In the case when the inertial sensors only consist of a single accelerometer, a new internal symmetry arises. As a consequence, the initial speed and orientation and the camera-inertial sensor transformation are not fully observable: all these quantities cannot be distinguished from the same quantities rotated around the accelerometer axis. All the remaining states are observable as in the standard visual-inertial sensor fusion problem.

It is very interesting to remark that these results provide a new insight about the problem of visual-vestibular integration for self-motion perception in neuroscience. Most of vertebrates are equipped with two distinct organs in the inner ear that are able to sense acceleration (both inertial acceleration and gravity). These are called otoliths and are the saccule and the utricle. The interesting point is that each of them is able to sense acceleration along two distinct axes. In accordance with our results, this makes possible to auto calibrate these sensors, separately. Specifically, our results clearly prove that there is enough information to estimate the position and orientation of a two-axis accelerometer with respect to the visual sensor, by only using the data provided by the two-axis accelerometer and the visual sensor. The auto calibration is a fundamental step to be accomplished in order to have the possibility of properly use the measurements provided by the sensor. In other words, if each otolith was constituted by a single axis accelerometer, the calibration was not possible. As a result, the measurements provided by this sensor were not properly used and the self-perception of motion was not possible.

Chapter 6

Analytic Derivations

In this chapter, we prove the validity of the analytic results presented in chapter 3. We start by introducing several important concepts starting by providing the definition of state indistinguishability in presence of unknown inputs 6.1. Section 6.2 is devoted to prove the results that hold in the case of a single unknown input ($m_w = 1$) and dynamics linear in the inputs ($g^0 = 0$). Then, in section 6.3, we extend these results to the general case.

6.1 State Indistinguishability, State augmentation and Observable Codistribution

In [18, 27] the observability properties of a nonlinear system driven by only known inputs are derived starting from the definition of indistinguishable states. According to this definition, it is proven the following fundamental property:

Proposition 1 *The Lie derivatives of any output computed along any direction allowed by the system dynamics take the same values at the states which are indistinguishable.*

Starting from this fundamental property it is possible to prove that algorithm 1 generates the observable codistribution [27]. In presence of unknown inputs, we first need to introduce a new definition of indistinguishability. To be conservative, the new definition must consider two states indistinguishable if there exists at least one pair of unknown inputs, such that, the outputs obtained starting from the first state under the effect of the first unknown input on a given time interval ($\mathcal{I}$), coincide with the outputs obtained starting from the second state under the effect of the second unknown input on the same time interval, and this holds for any choice of the known inputs.

We introduce the following definition:

Definition 2 (Indistinguishable states in presence of UI) *Two states x_a and x_b are indistinguishable if, for every $u(t)$ (the known input vector function), there exist $w_a(t)$ and $w_b(t)$ (i.e., two unknown input vector functions in general, but not necessarily, different from each other and at least one of them does not vanish) such that $h(x(t;\ x_a;\ u;\ w_a)) = h(x(t;\ x_b;\ u;\ w_b))$ $\forall t \in \mathcal{I}$.*

On the other hand, if the condition that the outputs coincide on $\mathcal{I}$ is achieved only for a unique pair of unknown inputs, the probability that this event occurs is zero. For this reason, we require that this condition is met for infinite pairs of unknown inputs. Unfortunately, the concept of measure in spaces with infinite dimensions is not trivial. In particular, there is no analogue of Lebesgue measure on an infinite-dimensional Banach space. One possibility, which is frequently adopted, is to use the concept of prevalent and shy sets [26]. Let us denote by $\mathcal{W}$ the functional space of all the possible unknown input functions. The probability that a given unknown input belongs to a shy subset of $\mathcal{W}$, is 0. The probability that a given unknown input belongs to a prevalent subset of $\mathcal{W}$, is 1. Finally, the probability that a given unknown input belongs to a non-shy subset of $\mathcal{W}$, is strictly larger than 0. In accordance with these remarks, we introduce the following definition:

Definition 3 (Indistinguishable states in presence of UI) *Two states x_a and x_b are indistinguishable if, for every $u(t)$ (the known input vector function), there exists a non-shy subset $\mathcal{W}_a$ in the functional space of all the possible unknown input functions, such that, for any unknown input function $w_a \in \mathcal{W}_a$ it exists an unknown input function w_b such that $h(x(t;\ x_a;\ u;\ w_a)) = h(x(t;\ x_b;\ u;\ w_b))$ $\forall t \in \mathcal{I}$.*

In the sequel we will adopt the second definition. In the case of driftless systems characterized by a single unknown input ($m_w = 1$), it is possible to obtain the same results by using the first definition.

The property stated by proposition 1 does not hold in presence of unknown inputs. Our first objective is to extend the original state in order to obtain a similar property for the resulting extended system. This new property will be the one stated by proposition 4.

To obtain such a result, we proceed as follows. We extend the original state by including the unknown inputs together with their time derivatives. Specifically, we denote by ${}^k x$ the extended state that includes the unknown inputs and their time derivatives up to the $(k-1)-$order:

$$ {}^k x \triangleq [x^T,\ w^T,\ w^{(1)\ T},\ \cdots,\ w^{(k-1)\ T}]^T \tag{6.1} $$

where $w^{(k)} \triangleq \frac{d^k w}{dt^k}$ and ${}^k x \in M^{(k)}$, with $M^{(k)}$ an open set of $\mathbb{R}^{n+km_w}$. From (2.1) it is immediate to obtain the dynamics for the extended state:

$$ {}^k \dot{x} = G + \sum_{i=1}^{m_u} F^i u_i + \sum_{j=1}^{m_w} W^j w_j^{(k)} \tag{6.2} $$

where:

$$ G \triangleq \begin{bmatrix} g^0 + \sum_{j=1}^{m_w} g^j w_j \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0_{m_w} \end{bmatrix} \tag{6.3} $$

$$ F^i \triangleq \begin{bmatrix} f^i \\ 0_{km_w} \end{bmatrix}, \quad W^j \triangleq \begin{bmatrix} 0_{n+(k-1)m_w+j-1} \\ 1 \\ 0_{m_w-j} \end{bmatrix} \tag{6.4} $$

and we denoted by 0_m the $m-$dimensional zero column vector. We remark that the resulting system has still m_u known inputs and m_w unknown inputs. However, while the m_u known inputs coincide with the original ones, the m_w unknown inputs are now the $k-$order time derivatives of the original unknown inputs. The state evolution depends on the known inputs via the vector fields F^i, $(i = 1, \cdots, m_u)$ and it depends on the unknown inputs via the unit vectors W^j, $(j = 1, \cdots, m_w)$. Finally, we remark that only the vector field G depends on the new state elements.

In the sequel, we will denote the extended system by $\Sigma^{(k)}$. Additionally, we use the notation: $\xi \triangleq [w^T,\ w^{(1)\ T},\ \cdots,\ w^{(k-1)\ T}]^T$. In this notation we have ${}^k x = [x^T,\ \xi^T]^T$. We also denote by $\Sigma^{(0)}$ the original system, i.e., the one characterized by the state x and the equations in (2.1).

We start by providing a simple result for $\Sigma^{(k)}$:

Lemma 1 *In $\Sigma^{(k)}$, the Lie derivatives of the output up to the m^{th} order ($m \le k$) are independent of $w_j^{(f)}$, $j = 1, \cdots, m_w$, $\forall f \ge m$.*

Proof: We proceed by induction on m for any k. When $m = 0$ we only have one zero-order Lie derivative (i.e., $h(x)$), which only depends on x, namely it is independent of $w^{(f)}$, $\forall f \ge 0$. Let us assume that the previous assert is true for m and let us prove that it holds for $m+1$. If it is true for m, any Lie derivative up to the m^{th} order is independent of $w^{(f)}$, for any $f \ge m$. In other words, the analytical expression of any Lie derivative up to the $m-$order is represented by a function $g(x, w, w^{(1)}, \cdots, w^{(m-1)})$. Hence, $\nabla g = [\frac{\partial g}{\partial x}, \frac{\partial g}{\partial w}, \frac{\partial g}{\partial w^{(1)}}, \cdots, \frac{\partial g}{\partial w^{(m-1)}}, 0_{(k-m)m_w}]$. It is immediate to realize that the product of this differential by any vector field in (6.2) depends at most on $w^{(m)}$, i.e., it is independent of $w^{(f)}$, $\forall f \ge m+1$ ■

A simple consequence of this lemma is the following property:

Proposition 2 *Let us consider the system $\Sigma^{(k)}$. The Lie derivatives of the output up to the k^{th} order along at least one vector among W^j $(j = 1, \cdots, m_w)$ are identically zero.*

Proof: From the previous lemma it follows that all the Lie derivatives, up to the $(k-1)-$order are independent of $w^{(k-1)}$, which are the last m_w components of the extended state in (6.1). Then, the proof follows from the fact that any vector among W^j $(j = 1, \cdots, m_w)$ has the first $n + (k-1)m_w$ components equal to zero ■

We have the following property:

Proposition 3 *The Lie derivatives of the output up to the k^{th} order along any vector field G, $F^1, \cdots, F^{m_u}$ for the system $\Sigma^{(k)}$ coincide with the same Lie derivatives for the system $\Sigma^{(k+1)}$*

Proof: We proceed by induction on m for any k. When $m = 0$ we only have one zero-order Lie derivative (i.e., $h(x)$), which is obviously the same for the two systems, $\Sigma^{(k)}$and $\Sigma^{(k+1)}$. Let us assume that the previous assert is true for m and let us prove that it holds for $m+1 \leq k$. If it is true for m, any Lie derivative up to the m^{th} order is the same for the two systems. Additionally, from lemma 1, we know that these Lie derivatives are independent of $w^{(f)}$, $\forall f \geq m$. The proof follows from the fact that the first $n + mm_w$ components of the vector fields G, $F^1, \cdots, F^{m_u}$ for $\Sigma^{(k)}$, coincide with the first $n + mm_w$ components of the same vector fields for $\Sigma^{(k+1)}$, when $m < k$ ■

For $\Sigma^{(k)}$ we have a fundamental property that is the extension of the one stated by proposition 1:

Proposition 4 *If x_a and x_b are indistinguishable, there exist ξ_a and ξ_b such that, in $\Sigma^{(k)}$, the Lie derivatives of the output up to the k^{th}-order, along all the vector fields that characterize the dynamics of $\Sigma^{(k)}$, take the same values at $[x_a, \xi_a]$ and $[x_b, \xi_b]$.*

Proof: We consider a piecewise-constant input $\tilde{u}$ as follows $(i = 1, \cdots, m_u)$:

$$\tilde{u}_i(t) = \tag{6.5}$$

$$\begin{cases} u_i^1 & t \in [0,\ t_1) \\ u_i^2 & t \in [t_1,\ t_1 + t_2) \\ \cdots & \\ u_i^g & t \in [t_1 + t_2 + \cdots + t_{g-1},\ t_1 + t_2 + \cdots + t_{g-1} + t_g) \end{cases}$$

Since x_a and x_b are indistinguishable, there exist two unknown input functions $w_a(t)$ and $w_b(t)$ such that the output coincide on x_a and x_b. In particular, we can write:

$$h(x(t;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)})) = h(x(t;\ [x_b, \xi_b];\ \tilde{u};\ w_b^{(k)})) \tag{6.6}$$

$\forall t \in [0,\ t_1 + t_2 + \cdots + t_{g-1} + t_g) \subset \mathcal{I}$. On the other hand, by taking the two quantities in (6.6) at $t = t_1 + t_2 + \cdots + t_{g-1} + t_g$, we can consider them as functions of the g arguments $t_1, t_2, \cdots, t_g$. Hence, by differentiating with respect to all these variables, we also have:

$$\frac{\partial^g h(x(t_1 + \cdots + t_g;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)}))}{\partial t_1 \partial t_2 \cdots \partial t_g} = \tag{6.7}$$

$$= \frac{\partial^g h(x(t_1 + \cdots + t_g;\ [x_b, \xi_b];\ \tilde{u};\ w_b^{(k)}))}{\partial t_1 \partial t_2 \cdots \partial t_g}$$

By computing the previous derivatives at $t_1 = t_2 = \cdots = t_g = 0$ and by using proposition 2 we obtain, if $g \le k$:

$$\mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h \Bigg|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h \Bigg|_{\substack{x = x_b \\ \xi = \xi_b}} \tag{6.8}$$

where $\theta_h = G + \sum_{i=1}^{m_u} F^i u^h_i$, $h = 1, \cdots, g$. The equality in (6.8) must hold for all possible choices of $u^h_1, \cdots, u^h_{m_u}$. By appropriately selecting these $u^h_1, \cdots, u^h_{m_u}$, we finally obtain:

$$\mathcal{L}^g_{v_1 v_2\cdots v_g} h \Bigg|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}^g_{v_1 v_2\cdots v_g} h \Bigg|_{\substack{x = x_b \\ \xi = \xi_b}} \tag{6.9}$$

where $v_1 v_2 \cdots v_g$ are vector fields belonging to the set $\{G, F^1, \cdots, F^{m_u}\}$ ■

By using the definition of shyness and the definition of indistinguishability previously introduced, it is possible to prove that, the statement of proposition 4, holds not only for a given pair ξ_a and ξ_b, but for infinite pairs. In particular, there exist a subset, U_a, of the Euclidean space $\mathbb{R}^{k+1}$, such that the property holds for any pair $\xi_a \in U_a$ and a given ξ_b. The fundamental point is that the subset U_a has measure larger than zero (in this case it is the measure of Lesbegues in the Euclidean space $\mathbb{R}^{k+1}$). Therefore, we can certainly find a pair ξ_a and ξ_b where the components of ξ_a are all different from zero. This property will be used in the sequel. We summarize this property with the following statement:

Remark 1 *The statement of proposition 4 holds for a pair ξ_a and ξ_b where all the components of ξ_a are different from zero.*

The main difference between propositions 1 and 4 is that, in the latter, we cannot consider any order Lie derivative, since the order cannot exceed k. This will have important consequences, as we will see. Before discussing this point, we remark that in [18] it was also defined the concept of $V-$indistinguishable states, with V a subset of the definition set that includes the two considered states. From this definition and the previous proof we can alleviate the assumptions in the previous proposition. Specifically, we have the following:

Remark 2 *The statement of proposition 4 also holds if x_a and x_b are $V-$indistinguishable.*

Now let us discuss how we can use the result stated by the proposition 4 to investigate the observability properties. Thanks to the results stated by propositions 3 and 4 we can easily build a codistribution that is observable. In the sequel, we will denote by $\mathcal{D}_E$ the differential with respect to the entire extended state. We will call *original state* the vector x and, we remind the reader that we denote by $\mathcal{D}$ the differential with respect to the original state. The observable codistribution for $\Sigma^{(k)}$ is the span of the differentials of all the Lie derivatives of the output along G, $F^1, \cdots, F^{m_u}$ up to the k-order. Hence, for any $m \le k$, it is obtained recursively by the following algorithm:

Algorithm 6 *Observable codistribution for the extended state ($m \le k$)*

1. $\bar{\Omega}_0 = span\{\mathcal{D}_E h\}$;

2. $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{F^i} \bar{\Omega}_{m-1}$

It is possible to obtain all the observability properties of the original state starting from the codistributions generated by the previous algorithm. In order to show this, we introduce the concept of *Observable Mode*. This concept can be also useful in the case without unknown inputs.
Given x_0, we denote by I_{x_0} the set of all the states x such that x and x_0 are indistinguishable. According to the theory of observability, a system is observable in x_0 if $I_{x_0} = x_0$. We introduce here the following new definition:

Definition 4 (Observable Mode) *A scalar function is observable in x_0 if it is constant on I_{x_0}. Additionally, it is weakly observable in x_0 if it exists an open neighbourhood B_{x_0} such that it is constant on $B_{x_0} \cap I_{x_0}$.*

The basic idea behind definition 4 is the following. Let us suppose that the true initial state is x_0. According to the definition of indistinguishable set (which is based on definition 3), the system only contains the information to establish whether the initial state belongs to the indistinguishable set I_{x_0} or not. Hence, if a scalar function takes the same value on this set, we conclude that the system has enough information to known the value of this scalar function at the initial time. We also remark that definition 4 generalizes the definition of observability. Specifically, a system is observable in x_0 when all the components of the state x are observable modes in x_0. When the scalar function is only weakly observable at x_0, we conclude that the system has enough information to know the value of this scalar function at the initial time, provided that we a priori know that the initial state is *sufficiently* close to x_0.

The following two propositions generalize theorem 3.1 and 3.11 in [18].

Proposition 5 *Given a scalar function $\theta(x)$, if it exists an integer k such that for the extended system $\Sigma^{(k)}$, $\mathcal{D}_E\theta \in \bar{\Omega}_k$ in a given x_0 and for a given extension ξ_0, then θ is weakly observable in x_0.*

Proof: $\theta(x)$ can be expressed in terms of the Lie derivatives of the output along the fields that characterize the dynamics of $\Sigma^{(k)}$, up to the $k-$order. We can write $\theta(x) = \mathcal{G}(\phi_1(x,\xi), \cdots, \phi_L(x,\xi))$, $\forall x \in B_{x_0}$ where $\phi_1(x,\xi), \cdots, \phi_L(x,\xi)$ are L Lie derivatives among the ones of above and $\mathcal{G}$ is a given function.

Let us consider a given x indistinguishable from x_0 and that belongs to B_{x_0}. In other words, $x \in B_{x_0} \cap I_{x_0}$. From proposition 4 there are two extensions ξ_a and ξ_b, such that, the Lie derivatives $\phi_1(x,\xi), \cdots, \phi_L(x,\xi)$ take the same values on the two extended states $[x, \xi_a]$ and $[x_0, \xi_b]$. Therefore, $\forall x \in B_{x_0} \cap I_{x_0}$ we have: $\theta(x) = \mathcal{G}(\phi_1(x,\xi_a), \cdots, \phi_L(x,\xi_a)) = \mathcal{G}(\phi_1(x_0,\xi_b), \cdots, \phi_L(x_0,\xi_b)) = \theta(x_0)$ and θ is weakly observable in x_0 $\blacksquare$

Proposition 6 *If the scalar function $\theta(x)$ is weakly observable in x_0, then $\exists k$ and ξ_0 such that $\mathcal{D}_E\theta \in \bar{\Omega}_k$ a.e. on an open neighbourhood B_{x_0}.*

Proof: We proceed by contradiction. $\forall k$ and ξ and for any open ball centered on x_0 with radius r ($B^r_{x_0}$), it exists a set $C^r_{x_0} \subseteq B^r_{x_0}$ with measure strictly larger than zero for which $\mathcal{D}_E\theta \notin \bar{\Omega}_k$. This means that $\forall k$ and ξ it exists a distribution with dimension larger than zero that is orthogonal to the codistribution $\bar{\Omega}_k$ in $C^r_{x_0}$. This means that it exists at least one vector field s, of the same dimension of the original state, such that, $\forall k$ and ξ, it exists at least one vector field s_k of dimension km_w, such that the vector field $[s^T, s_k^T]^T$ is orthogonal to the codistribution $\bar{\Omega}_k$ in $C^r_{x_0}$. From this, we obtain that it exists $\epsilon > 0$ such that the states x_0 and $x_0 + \epsilon s$ are indistinguishable. Indeed, it is possible to express the $m-$time derivative of the output at the

initial time, for a generic integer m, in terms of the Lie derivatives of the output along the fields of $\Sigma^{(k)}$, up to the $m-$order, and the known inputs and their time derivatives (see [39], sect 2.5.1). Since this holds for any k, the $m-$order time derivative of the output at the initial time coincides in x_0 and $x_0+\epsilon s$. Since this holds for any order m, from the Taylor theorem the output coincides on a given time interval and consequently x_0 and $x_0+\epsilon s$ are indistinguishable. But this means that also $\theta(x_0) = \theta(x_0+\epsilon s)$. Hence, $\mathcal{D}_E\theta \cdot s = 0$ a.e. on an open neighbourhood B_{x_0} and $\mathcal{D}_E\theta \in \bar{\Omega}_k$ a.e. on an open neighbourhood B_{x_0} $\blacksquare$

Propositions 5 and 6 state that all the observability properties of the original state are contained in the codistributions generated by algorithm 6. From this, we can easily obtain a sufficient condition for the observability of the original state. Indeed, if on a given x_0, the differential of a given component of x (the original state) belongs to $\bar{\Omega}_m$ for a given integer $m \leq k$, we can conclude that this state component is weakly locally observable (in x_0). If this holds for all the state components, we can conclude that the entire original state is weakly locally observable. More in general, we can conclude that, a given scalar function of the original state, is weakly observable in a given point x_0 if its differential (computed in x_0) belongs to $\bar{\Omega}_m$ (computed in x_0) for a given integer $m \leq k$. On the other hand, we remark the following two fundamental differences between algorithms 1 and 6:

1. In the latter, since the state augmentation can be continued indefinitely, we do not have convergence;

2. The latter provides a codistribution that describes simultaneously the observability properties of the original state and its extension.

The goal of the next two sections is to address these fundamental issues. In particular, we show that it is possible to directly compute the entire observable codistribution of the original system, namely, without the need of extending the state. In section 6.2 we start by dealing with the case of a single unknown input ($m_w = 1$) and dynamics linear in the inputs ($g^0 = 0$). In this case the entire observable codistribution of the original system (i.e., without extension) is the one computed by algorithm 3. In section 6.3 we consider the general case (i.e., $\forall m_w$ and dynamics affine in the inputs). In this case the entire observable codistribution of the original system is the one computed by algorithm 5. For both algorithms 3 and 5, we analytically derive the convergence criteria, which are the ones provided in chapter 3.

6.2 Proof of the validity of the analytic criterion in the driftless case and with a single unknown input

This section is devoted to the case of a single unknown input (i.e., $m_w = 1$) and dynamics linear in the inputs ($g^0 = 0$). In other words, we are considering the system characterized by (2.2). When $m_w = 1$, the extended state that includes the time derivatives of w up to the $(k-1)-$order is:

$$^k x \triangleq [x^T, \ w, \ w^{(1)}, \ \cdots , \ w^{(k-1)}]^T \tag{6.10}$$

where $w^{(j)} \triangleq \frac{d^j w}{dt^j}$. The dimension of the extended state is in this case $n + k$. For the clarity sake, let us consider the case of a single known input (i.e., we consider the system in (2.2) with $m_u = 1$). We provide all the analytic results in this simpler case and then we extend them to the case of $m_u > 1$. From (2.1) it is immediate to obtain the dynamics for the extended state:

$$^k \dot{x} = G(^k x) + F(x)u + W w^{(k)} \tag{6.11}$$

where:

$$F \triangleq \begin{bmatrix} f(x) \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{bmatrix} \quad G \triangleq \begin{bmatrix} g(x)w \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0 \end{bmatrix} \quad W \triangleq \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \\ 1 \end{bmatrix} \tag{6.12}$$

and we set $f(x) \triangleq f^1(x)$.

Algorithm 6 becomes:

Algorithm 7 *Observable codistribution for the extended state in the case $m_u = 1$ ($m \leq k$)*

1. $\bar{\Omega}_0 = span\{\mathcal{D}_E h\}$;
2. $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1}$

where G is now the simpler vector field given in (6.12) and $\mathcal{D}_E$ denotes the differential respect to the state in (6.10).

We will address the two fundamental issues mentioned at the end of the previous section. This is obtained into two separates steps. In the first step (sect. 6.2.1) we perform a separation on the observable codistribution defined by algorithm 7. This codistribution can be split into two codistributions: the former is the codistribution generated by algorithm 3, once embedded in the extended space, and the latter is the codistribution L^m defined in section 6.2.1 (see theorem 1). We prove (Lemma 2) that the second codistribution (L^m) can be ignored when deriving the observability properties of the original state. In the second step (sect. 6.2.2) we prove that algorithm 3 converges in at most $n + 2$ steps and we provide the convergence criterion. Finally, in section 6.2.3 we extend the results of the previous two theorems to the case of multiple known inputs ($m_u > 1$).

6.2.1 Separation

For each integer m, we generate the codistribution Ω_m by using algorithm 3 (note that here we are considering $m_u = 1$). By construction, the generators of Ω_m are the differentials of scalar functions that only depend on the original state (x) and not on its extension. In the sequel, we need to embed this codistribution in $\mathbb{R}^{n+k}$. We will denote by $[\Omega_m, 0_k]$ the codistribution made by covectors whose first n components are covectors in Ω_m and the last components are all zero. Additionally, we will denote by L^m the codistribution that is the span of the Lie derivatives of $\mathcal{D}_E h$ up to the order m along the vector G, i.e., $L^m \triangleq span\{\mathcal{L}_G^1 \mathcal{D}_E h, \mathcal{L}_G^2 \mathcal{D}_E h, \cdots, \mathcal{L}_G^m \mathcal{D}_E h\}$. We finally introduce the following codistribution:

Definition 5 ($\tilde{\Omega}$ codistribution) *This codistribution is defined as follows:* $\tilde{\Omega}_m \triangleq [\Omega_m, 0_k] + L^m$

The codistribution $\tilde{\Omega}_m$ consists of two parts. Specifically, we can select a basis that consists of exact differentials that are the differentials of functions that only depend on the original state (x) and not on its extension (these are the generators of $[\Omega_m, 0_k]$) and the differentials $\mathcal{L}_G^1 \mathcal{D}_E h, \mathcal{L}_G^2 \mathcal{D}_E h, \cdots, \mathcal{L}_G^m \mathcal{D}_E h$. The second set of generators, i.e., the differentials $\mathcal{L}_G^1 \mathcal{D}_E h, \mathcal{L}_G^2 \mathcal{D}_E h, \cdots, \mathcal{L}_G^m \mathcal{D}_E h$, are m and, with respect to the first set, they are differentials of functions that also depend on the state extension $\xi = [w,\ w^{(1)}, \cdots,\ w^{(m-1)}]^T$. We have the following result:

Lemma 2 *Let us denote with x_j the j^{th} component of the state ($j = 1, \cdots, n$). We have: $\mathcal{D}x_j \in \Omega_m$ if and only if $\mathcal{D}_E x_j \in \tilde{\Omega}_m$*

Proof: The fact that $\mathcal{D}x_j \in \Omega_m$ implies that $\mathcal{D}_E x_j \in \tilde{\Omega}_m$ is obvious since $[\Omega_m, 0_k] \subseteq \tilde{\Omega}_m$ by definition. Let us prove that also the contrary holds, i.e., that if $\mathcal{D}_E x_j \in \tilde{\Omega}_m$ then $\mathcal{D}x_j \in \Omega_m$. Since $\mathcal{D}_E x_j \in \tilde{\Omega}_m$ we have $\mathcal{D}_E x_j = \sum_{i=1}^{N_1} c_i^1 \omega_i^1 + \sum_{i=1}^{N_2} c_i^2 \omega_i^2$, where $\omega_1^1, \omega_2^1, \cdots, \omega_{N_1}^1$ are N_1 generators of $[\Omega_m, 0_k]$, $\omega_1^2, \omega_2^2, \cdots, \omega_{N_2}^2$ are N_2 generators of L^m and $c_1^1, \cdots, c_{N_1}^1, c_1^2, \cdots, c_{N_2}^2$ are suitable coefficients. We want to prove that $N_2 = 0$.

We proceed by contradiction. Let us suppose that $N_2 \geq 1$. We remark that the first set of generators have the last k entries equal to zero, as for $\mathcal{D}_E x_j$. The second set of generators consists of the Lie derivatives of $\mathcal{D}_E h$ along G up to the m order. Let us select the one that is the highest order Lie derivative and let us denote by j' this highest order. We have $1 \leq N_2 \leq j' \leq m$. By a direct computation, it is immediate to realize that this is the only generator that depends on $w^{(j'-1)}$. Specifically, the dependence is linear by the product $L_g^1 w^{(j'-1)}$ (we remind the reader that $L_g^1 \neq 0$). But this means that $\mathcal{D}_E x_j$ has the $(n+j')^{th}$ entry equal to $L_g^1 \neq 0$ and this is not possible since $\mathcal{D}_E x_j = [\mathcal{D}x_j, 0_k]$ $\blacksquare$

A fundamental consequence of this lemma is that, if we are able to prove that $\tilde{\Omega}_m = \bar{\Omega}_m$, the weak local observability of the original state x, can be investigated by only considering the codistribution Ω_m. In the rest of this section we prove this fundamental theorem, stating that $\tilde{\Omega}_m = \bar{\Omega}_m$.

For a given $m \leq k$ we define the vector $\Phi_m \in \mathbb{R}^{n+k}$ by the following algorithm:

1. $\Phi_0 = F$;
2. $\Phi_m = [\Phi_{m-1},\ G]$

where now the Lie brackets $[\cdot, \cdot]$ are computed with respect to the extended state, i.e.:

$$[F,\ G] \triangleq \frac{\partial G}{\partial^k x} F(^k x) - \frac{\partial F}{\partial^k x} G(^k x)$$

By a direct computation it is easy to realize that Φ_m has the last k components identically null. In the sequel, we will denote by $\breve{\Phi}_m$ the vector in $\mathbb{R}^n$ that contains the first n components of Φ_m. In other words, $\Phi_m \triangleq [\breve{\Phi}_m^T, 0_k^T]^T$. Additionally, we set $\hat{\phi}_m \triangleq \begin{bmatrix} \phi_m \\ 0_k \end{bmatrix}$ (ϕ_m is defined by algorithm 2).

We have the following result:

Lemma 3 $\mathcal{L}_G\bar{\Omega}_m + \mathcal{L}_{\Phi_m}\mathcal{D}_E h = \mathcal{L}_G\bar{\Omega}_m + \mathcal{L}_F\mathcal{L}_G^m\mathcal{D}_E h$

Proof: We have $\mathcal{L}_F\mathcal{L}_G^m\mathcal{D}_E h = \mathcal{L}_G\mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h + \mathcal{L}_{\Phi_1}\mathcal{L}_G^{m-1}\mathcal{D}_E h$.

The first term $\mathcal{L}_G\mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h \in \mathcal{L}_G\bar{\Omega}_m$. Hence, we need to prove that $\mathcal{L}_G\bar{\Omega}_m + \mathcal{L}_{\Phi_m}\mathcal{D}_E h = \mathcal{L}_G\bar{\Omega}_m + \mathcal{L}_{\Phi_1}\mathcal{L}_G^{m-1}\mathcal{D}_E h$. We repeat the previous procedure m times. Specifically, we use the equality $\mathcal{L}_{\Phi_j}\mathcal{L}_G^{m-j}\mathcal{D}_E h = \mathcal{L}_G\mathcal{L}_{\Phi_j}\mathcal{L}_G^{m-j-1}\mathcal{D}_E h + \mathcal{L}_{\Phi_{j+1}}\mathcal{L}_G^{m-j-1}\mathcal{D}_E h$, for $j = 1, \cdots, m$, and we remove the first term since $\mathcal{L}_G\mathcal{L}_{\Phi_j}\mathcal{L}_G^{m-j-1}\mathcal{D}_E h \in \mathcal{L}_G\bar{\Omega}_m$ ■

Lemma 4 $\breve{\Phi}_m = \sum_{j=1}^m c_j^n(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\phi_j$, *i.e., the vector* $\breve{\Phi}_m$ *is a linear combination of the vectors* ϕ_j *(*$j = 1, \cdots, m$*), where the coefficients (*c_j^n*) depend on the state only through the functions that generate the codistribution* L^m

Proof: We proceed by induction. By a direct computation it is immediate to obtain: $\breve{\Phi}_1 = \phi_1\mathcal{L}_G h$.
Inductive step: Let us assume that $\breve{\Phi}_{m-1} = \sum_{j=1}^{m-1} c_j(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h)\phi_j$. We have:

$$\Phi_m = [\Phi_{m-1},\ G] = \sum_{j=1}^{m-1}\left[c_j \begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G\right] =$$

$$\sum_{j=1}^{m-1} c_j \left[\begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G\right] - \sum_{j=1}^{m-1} \mathcal{L}_G c_j \begin{bmatrix} \phi_j \\ 0_k \end{bmatrix}$$

We directly compute the Lie bracket in the sum (note that ϕ_j is independent of the unknown input w and its time derivatives):

$$\left[\begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G\right] = \begin{bmatrix} [\phi_j,\ g]w \\ 0_k \end{bmatrix} = \begin{bmatrix} \phi_{j+1}\mathcal{L}_G^1 h \\ 0_k \end{bmatrix}$$

Regarding the second term, we remark that $\mathcal{L}_G c_j = \sum_{i=1}^{m-1} \frac{\partial c_j}{\partial(\mathcal{L}_G^i h)}\mathcal{L}_G^{i+1} h$. By setting $\tilde{c}_j = c_{j-1}\mathcal{L}_G^1 h$ for $j = 2, \cdots, m$ and $\tilde{c}_1 = 0$, and by setting $\bar{c}_j = -\sum_{i=1}^{m-1} \frac{\partial c_j}{\partial(\mathcal{L}_G^i h)}\mathcal{L}_G^{i+1} h$ for $j = 1, \cdots, m-1$ and $\bar{c}_m = 0$, we obtain $\breve{\Phi}_m = \sum_{j=1}^m (\tilde{c}_j + \bar{c}_j)\phi_j$, which proves our assert since $c_j^n (\triangleq \tilde{c}_j + \bar{c}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h$ ■

It also holds the following result:

Lemma 5 *If* $w \neq 0$, $\hat{\phi}_m = \sum_{j=1}^m b_j^n(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\Phi_j$, *i.e., the vector* $\hat{\phi}_m$ *is a linear combination of the vectors* Φ_j *(*$j = 1, \cdots, m$*), where the coefficients (*b_j^n*) depend on the state only through the functions that generate the codistribution* L^m

Proof: We proceed by induction. By a direct computation it is immediate to obtain: $\hat{\phi}_1 = \Phi_1 \frac{1}{\mathcal{L}_G h}$ (note that $\mathcal{L}_G h = L_g^1 w \neq 0$).
Inductive step: Let us assume that $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h)\Phi_j$. We need to prove that $\hat{\phi}_m = \sum_{j=1}^m b_j^n(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\Phi_j$. We start by applying on both members of

the equality $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h)\Phi_j$ the Lie bracket with respect to G. We obtain for the first member: $[\hat{\phi}_{m-1},\ G] = \hat{\phi}_m \mathcal{L}_G^1 h$. For the second member we have:

$$\sum_{j=1}^{m-1}[b_j\Phi_j,\ G] = \sum_{j=1}^{m-1} b_j[\Phi_j,\ G] - \sum_{j=1}^{m-1}\mathcal{L}_G b_j \Phi_j =$$

$$= \sum_{j=1}^{m-1} b_j\Phi_{j+1} - \sum_{j=1}^{m-1}\sum_{i=1}^{m-1}\frac{\partial b_j}{\partial(\mathcal{L}_G^i h)}\mathcal{L}_G^{i+1}h\Phi_j$$

Since $\mathcal{L}_G h = L_g^1 w \neq 0$, by setting $\tilde{b}_j = \frac{b_{j-1}}{\mathcal{L}_G^1 h}$ for $j = 2, \cdots, m$ and $\tilde{b}_1 = 0$, and by setting $\bar{b}_j = -\sum_{i=1}^{m-1}\frac{\partial b_j}{\partial(\mathcal{L}_G^i h)}\frac{\mathcal{L}_G^{i+1}h}{\mathcal{L}_G^1 h}$ for $j = 1, \cdots, m-1$ and $\bar{b}_m = 0$, we obtain $\hat{\phi}_m = \sum_{j=1}^{m}(\tilde{b}_j + \bar{b}_j)\Phi_j$, which proves our assert since $b_j^n (\triangleq \tilde{b}_j + \bar{b}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h$ $\blacksquare$

An important consequence of the previous two lemmas is the following result:

Proposition 7 *If $w \neq 0$, the following two codistributions coincide:*

1. *span*$\{\mathcal{L}_{\Phi_0}\mathcal{D}_E h, \mathcal{L}_{\Phi_1}\mathcal{D}_E h, \cdots, \mathcal{L}_{\Phi_m}\mathcal{D}_E h, \mathcal{L}_G^1\mathcal{D}_E h, \cdots \mathcal{L}_G^m\mathcal{D}_E h\}$;
2. *span*$\{\mathcal{L}_{\hat{\phi}_0}\mathcal{D}_E h, \mathcal{L}_{\hat{\phi}_1}\mathcal{D}_E h, \cdots, \mathcal{L}_{\hat{\phi}_m}\mathcal{D}_E h, \mathcal{L}_G^1\mathcal{D}_E h, \cdots \mathcal{L}_G^m\mathcal{D}_E h\}$;

We are now ready to prove the following fundamental result:

Theorem 1 (Separation) *If $w \neq 0$, $\bar{\Omega}_m = \tilde{\Omega}_m \triangleq [\Omega_m, 0_k] + L^m$*

Proof: We proceed by induction. By definition, $\bar{\Omega}_0 = \tilde{\Omega}_0$ since they are both the span of $\mathcal{D}_E h$.
Inductive step: Let us assume that $\bar{\Omega}_{m-1} = \tilde{\Omega}_{m-1}$. We have: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_F\bar{\Omega}_{m-1} + \mathcal{L}_G\bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + \mathcal{L}_F\tilde{\Omega}_{m-1} + \mathcal{L}_G\bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_F L^{m-1} + \mathcal{L}_G\bar{\Omega}_{m-1}$. On the other hand, $\mathcal{L}_F L^{m-1} = \mathcal{L}_F\mathcal{L}_G^1\mathcal{D}_E h + \cdots + \mathcal{L}_F\mathcal{L}_G^{m-2}\mathcal{D}_E h + \mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h$. The first $m-2$ terms are in $\bar{\Omega}_{m-1}$. Hence we have: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h + \mathcal{L}_G\bar{\Omega}_{m-1}$. By using lemma 3 we obtain: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h + \mathcal{L}_G\bar{\Omega}_{m-1}$. By using again the induction assumption we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h + \mathcal{L}_G[\Omega_{m-1}, 0_k] + \mathcal{L}_G L^{m-1} = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h + [\mathcal{L}_{\frac{g}{L_g^1}}\Omega_{m-1}, 0_k]$, where we used $L^m + \mathcal{L}_G[\Omega_{m-1}, 0_k] = L^m + [\mathcal{L}_{\frac{g}{L_g^1}}\Omega_{m-1}, 0_k]$, which holds because $\mathcal{L}_G h = L_g^1 w \neq 0$. By using proposition 7, we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f\Omega_{m-1}, 0_k] + \mathcal{L}_{\hat{\phi}_{m-1}}\mathcal{D}_E h + [\mathcal{L}_{\frac{g}{L_g^1}}\Omega_{m-1}, 0_k] = \tilde{\Omega}_m$ $\blacksquare$

Theorem 1 is fundamental. It allows us to obtain all the observability properties of the original state by restricting the computation to the codistribution defined by algorithm 3, namely a codistribution whose covectors have the same dimension of the original space. In other words, the dimension of these covectors is independent of the state augmentation.

6.2.2 Convergence

Algorithm 3 is recursive and $\Omega_m \subseteq \Omega_{m+1}$. This means that, if for a given m the differentials of the components of the original state belong to Ω_m, we can conclude that the original state is weakly locally observable. On the other hand, if this is not true, we cannot exclude that it is true for a larger m. The goal of this section is precisely to address this issue. We will show

that the algorithm converges in a finite number of steps and we will also provide the criterion to establish that the algorithm has converged (theorem 2). This theorem will be proved at the end of this section since we need to introduce several important new quantities and properties.
When investigating the convergence properties of algorithm 3, we remark that, the main difference between algorithm 1 and 3, is the presence of the last term in the recursive step of the latter. Without this term, the convergence criterion would simply consist of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1.

The following result provides the convergence criterion in a very special case that basically occurs when the contribution due to the last term in the recursive step of algorithm 3 is included in the other terms. In this case, we obviously obtain that the convergence criterion consists of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1. For any integer $j \geq 0$ we define:

$$\chi_j \triangleq \frac{\mathcal{L}_{\phi_j} L^1_g}{L^1_g} \tag{6.13}$$

We have the following result:

Lemma 6 *Let us denote by Λ_j the distribution generated by $\phi_0, \phi_1, \cdots, \phi_j$ and by $m(\leq n-1)$ the smallest integer for which $\Lambda_{m+1} = \Lambda_m$ (n is the dimension of the state x). In the very special case when $\chi_j = 0$, $\forall j = 0, \cdots, m$, algorithm 3 converges at the integer j such that $\Omega_{j+1} = \Omega_j$ and this occurs in at most $n-1$ steps.*

Proof: First of all, we remind the reader that the existence of an integer $m(\leq n-1)$ such that $\Lambda_{m+1} = \Lambda_m$ is proved in [27]. In particular, the first chapter in [27] analyzes the convergence of Λ_j with respect to j. It is proved that the distribution converges to Λ^* and that the convergence is achieved at the smallest integer for which we have $\Lambda_{m+1} = \Lambda_m$. Additionally, we have $\Lambda_{m+1} = \Lambda_m = \Lambda^*$ and m cannot exceed $n-1$.

In the very special case when $\chi_j = 0$, $\forall j = 0, \cdots, m$, thanks to the aforementioned convergence of the distribution Λ_j, we easily obtain that $\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h = 0$ $\forall j \geq 1$. Now, let us consider the following equation:

$$\mathcal{L}_{\phi_j} h = \frac{1}{L^1_g} \left(\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h - \mathcal{L}_g \mathcal{L}_{\phi_{j-1}} h \right) \tag{6.14}$$

Since $\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h = 0$ $\forall j \geq 1$, we have $\mathcal{L}_{\phi_j} h = -\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{L}_{\phi_{j-1}} h$, for any $j \geq 1$. Therefore, we conclude that, the last term in the recursive step of algorithm 3, is included in the second last term and, in this special case, algorithm 3 has converged when $\Omega_{m+1} = \Omega_m$. This occurs in at most $n-1$ steps, as for algorithm 1. ■
Let us consider now the general case. To proceed we need to introduce several important new quantities and properties.
For a given positive integer j we define the vector $\psi_j \in \mathbb{R}^n$ by the following algorithm:

1. $\psi_0 = f$;
2. $\psi_j = [\psi_{j-1}, \ \frac{g}{L^1_g}]$

It is possible to find the expression that relates these vectors to the vectors ϕ_j, previously defined. Specifically we have:

Lemma 7 *It holds the following equation:*

$$\psi_j = \phi_j + \left\{ \sum_{i=0}^{j-1} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1} \tag{6.15}$$

Proof: We proceed by induction. By definition $\psi_0 = \phi_0 = f$ and equation (6.15) holds for $j = 0$.
Inductive step: Let us assume that it holds for a given $j - 1 \geq 0$ and let us prove its validity for j. We have:

$$\psi_j = \left[\psi_{j-1}, \ \frac{g}{L_g^1} \right] = \left[\phi_{j-1}, \ \frac{g}{L_g^1} \right]$$

$$+ \left[\left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}, \ \frac{g}{L_g^1} \right]$$

On the other hand:

$$\left[\phi_{j-1}, \ \frac{g}{L_g^1} \right] = \phi_j - \frac{\mathcal{L}_{\phi_{j-1}} L_g^1}{L_g^1} \frac{g}{L_g^1}$$

and

$$\left[\left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}, \ \frac{g}{L_g^1} \right] =$$

$$-\mathcal{L}_{\frac{g}{L_g^1}} \left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1} =$$

$$\left\{ \sum_{i=0}^{j-2} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}$$

Hence:

$$\psi_j = \phi_j - \frac{\mathcal{L}_{\phi_{j-1}} L_g^1}{L_g^1} \frac{g}{L_g^1} + \left\{ \sum_{i=0}^{j-2} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1},$$

which coincides with (6.15) ■

We have the following result:

Lemma 8 *For $i = 0, 1, \cdots, m-2$, we have:*

$$\mathcal{D} \frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \in \Omega_m \tag{6.16}$$

Proof: By construction, $\mathcal{D}\mathcal{L}_{\phi_i} h \in \Omega_m$, for any $i = 1, \cdots, m-1$. On the other hand, we have:

$$\mathcal{L}_{\phi_i} h = \frac{1}{L_g^1} [\mathcal{L}_{\phi_{i-1}} \mathcal{L}_g h - \mathcal{L}_g \mathcal{L}_{\phi_{i-1}} h] = \frac{\mathcal{L}_{\phi_{i-1}} L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}} \mathcal{L}_{\phi_{i-1}} h$$

We compute the differential of both members of this equation. Since $\mathcal{D}\mathcal{L}_{\frac{g}{L_g^1}}\mathcal{L}_{\phi_{i-1}}h \in \Omega_m$, for any $i = 1, \cdots, m-1$, also $\mathcal{D}\frac{\mathcal{L}_{\phi_{i-1}}L_g^1}{L_g^1} \in \Omega_m$ ■

From lemma 7 with $j = 1, \cdots, m-1$ and lemma 8 it is immediate to obtain the following result:

Proposition 8 *If Ω_m is invariant with respect to $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$ then it is also invariant with respect to $\mathcal{L}_{\phi_j}$, $j = 1, \cdots, m-1$.*

In order to obtain the convergence criterion for algorithm 3 we need to substitute the expression of ϕ_j in terms of ϕ_{j-2} in the term $\mathcal{L}_{\phi_j}h$. This will allow us to detect the key quantity that governs the convergence of algorithm 3, in particular regarding the contribution due to the last term in the recursive step. In the case $m_w = 1$, $g^0 = 0$, this quantity is a scalar and it is the one provided in (3.2). For the sake of clarity, we provide equation (3.2) below:

$$\tau \triangleq \frac{L_g^2}{(L_g^1)^2}$$

where $L_g^2 \triangleq \mathcal{L}_g^2 h$. The behaviour of the last term in the recursive step of algorithm 3 is given by the following lemma:

Lemma 9 *We have the following key equality:*

$$\mathcal{L}_{\phi_j}h = \mathcal{L}_{\phi_{j-2}}\tau + \tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}}\left(\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} + \mathcal{L}_{\phi_{j-1}}h\right) \tag{6.17}$$

$j \geq 2$.

Proof: We will prove this equality by an explicit computation. We have:

$$\mathcal{L}_{\phi_j}h = \frac{1}{L_g^1}\left(\mathcal{L}_{\phi_{j-1}}\mathcal{L}_g h - \mathcal{L}_g\mathcal{L}_{\phi_{j-1}}h\right)$$

The second term on the right hand side simplifies with the last term in (6.17). Hence we have to prove:

$$\frac{1}{L_g^1}\mathcal{L}_{\phi_{j-1}}L_g^1 = \mathcal{L}_{\phi_{j-2}}\tau + \tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}}\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} \tag{6.18}$$

We have:

$$\frac{1}{L_g^1}\mathcal{L}_{\phi_{j-1}}L_g^1 = \frac{1}{(L_g^1)^2}\left(\mathcal{L}_{\phi_{j-2}}L_g^2 - \mathcal{L}_g\mathcal{L}_{\phi_{j-2}}L_g^1\right) \tag{6.19}$$

We remark that:

$$\frac{1}{(L_g^1)^2}\mathcal{L}_{\phi_{j-2}}L_g^2 = \mathcal{L}_{\phi_{j-2}}\tau + 2\tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1}$$

and

$$\frac{1}{(L_g^1)^2}\mathcal{L}_g\mathcal{L}_{\phi_{j-2}}L_g^1 = \tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} + \mathcal{L}_{\frac{g}{L_g^1}}\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1}$$

By substituting these two last equalities in (6.19) we immediately obtain (6.18) ■

Lemma 10 *In general, it exists an integer $m \leq n+2$ (being n the dimension of x) such that $\mathcal{D}\tau \in \Omega_m$.*

Proof: Let us introduce the following notation, for a given integer j:

- $\mathcal{Z}_j \triangleq \mathcal{L}_{\phi_{j+2}} h$;
- $\mathcal{B}_j \triangleq \mathcal{L}_{\phi_j} \tau$;
- $\chi_j \triangleq \frac{\mathcal{L}_{\phi_j} L^1_g}{L^1_g}$.

By construction, $\mathcal{D}\mathcal{Z}_j \in \Omega_{j+3}$. On the other hand, from equation (6.17), we immediately obtain:

$$\mathcal{D}\mathcal{Z}_j = \mathcal{D}\mathcal{B}_j + \chi_j \mathcal{D}\tau + \tau \mathcal{D}\chi_j - \mathcal{L}_{\frac{g}{L^1_g}} \left(\mathcal{D}\chi_j + \mathcal{D}\mathcal{L}_{\phi_{j+1}} h \right) \tag{6.20}$$

By using lemma 8 we obtain the following results:

- $\tau \mathcal{D}\chi_j \in \Omega_{j+2}$;
- $\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{D}\chi_j \in \Omega_{j+3}$.

Additionally, $\mathcal{L}_{\frac{g}{L^1_g}} \mathcal{D}\mathcal{L}_{\phi_{j+1}} h \in \Omega_{j+3}$. Hence, from (6.20), we obtain that the following covector:

$$\mathcal{Z}'_j \triangleq \mathcal{D}\mathcal{B}_j + \chi_j \mathcal{D}\tau \tag{6.21}$$

belongs to Ω_{j+3}. Let us denote by j^* the smallest integer such that:

$$\mathcal{D}\mathcal{B}_{j^*} = \sum_{j=0}^{j^*-1} c_j \mathcal{D}\mathcal{B}_j + c_{-1}\mathcal{D}h \tag{6.22}$$

Note that j^* is a finite integer and in particular $j^* \leq n-1$. Indeed, if this would not be the case, the dimension of the codistribution generated by $\mathcal{D}h, \mathcal{D}\mathcal{B}_0, \mathcal{D}\mathcal{B}_1, \cdots, \mathcal{D}\mathcal{B}_{n-1}$ would be $n+1$, i.e., larger than n. From (6.22) and (6.21) we obtain:

$$\mathcal{Z}'_{j*} = \sum_{j=0}^{j^*-1} c_j \mathcal{D}\mathcal{B}_j + c_{-1}\mathcal{D}h + \chi_{j*}\mathcal{D}\tau \tag{6.23}$$

From equation (6.21), for $j = 0, \cdots, j^*-1$, we obtain: $\mathcal{D}\mathcal{B}_j = \mathcal{Z}'_j - \chi_j \mathcal{D}\tau$. By substituting in (6.23) we obtain:

$$\mathcal{Z}'_{j*} - \sum_{j=0}^{j^*-1} c_j \mathcal{Z}'_j - c_{-1}\mathcal{D}h = \left(- \sum_{j=0}^{j^*-1} c_j \chi_j + \chi_{j*} \right) \mathcal{D}\tau \tag{6.24}$$

We remark that the left hand side consists of the sum of covectors that belong to Ω_{j^*+3}. Since in general $\chi_{j*} \neq \sum_{j=0}^{j^*-1} c_j \chi_j$, we have $\mathcal{D}\tau \in \Omega_{j^*+3}$. By setting $m \triangleq j^*+3$, we have $m \leq n+2$ and $\mathcal{D}\tau \in \Omega_m$ ■

The previous lemma ensures that, in general, it exists a finite $m \leq n+2$ such that $\mathcal{D}\tau \in \Omega_m$. Note that the previous proof holds if the quantity $\chi_{j*} - \sum_{j=0}^{j^*-1} c_j \chi_j$ does not vanish. This holds in general, with the exception of the trivial case considered in lemma 6, in which case $\chi_j = 0$, $\forall j$.

The following theorem allows us to obtain the criterion to stop algorithm 3:

Theorem 2 *If $\mathcal{D}\tau \in \Omega_m$ and Ω_m is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$, then $\Omega_{m+p} = \Omega_m$ $\forall p \geq 0$*

Proof: We proceed by induction. Obviously, the equality holds for $p = 0$.
Inductive step: let us assume that $\Omega_{m+p} = \Omega_m$ and let us prove that $\Omega_{m+p+1} = \Omega_m$. We have to prove that $\mathcal{D}\mathcal{L}_{\phi_{m+p}}h \in \Omega_m$. Indeed, from the inductive assumption, we know that $\Omega_{m+p}(= \Omega_m)$ is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$. Additionally, because of this invariance, by using proposition 8, we obtain that Ω_m is also invariant under $\mathcal{L}_{\phi_j}$, for $j = 1, 2, \cdots, m+p-1$. Since $\mathcal{D}\tau \in \Omega_m$ we have $\mathcal{D}\mathcal{L}_{\phi_{m+p-2}}\tau \in \Omega_m$. Additionally, $\mathcal{D}\mathcal{L}_{\phi_{m+p-1}}h \in \Omega_m$ and, because of lemma 8, we also have $\mathcal{D}\frac{\mathcal{L}_{\phi_{m+p-2}}L_g^1}{L_g^1} \in \Omega_m$. Finally, because of the invariance under $\mathcal{L}_{\frac{g}{L_g^1}}$, also the Lie derivatives along $\frac{g}{L_g^1}$ of $\mathcal{D}\mathcal{L}_{\phi_{m+p-1}}h$ and $\mathcal{D}\frac{\mathcal{L}_{\phi_{m+p-2}}L_g^1}{L_g^1}$ belong to Ω_m. Now, we use equation (6.17) for $j = m+p$. By computing the differential of this equation it is immediate to obtain that $\mathcal{D}\mathcal{L}_{\phi_{m+p}}h \in \Omega_m$ ■

We conclude this section by providing an upper bound for the number of steps that are in general necessary to achieve the convergence. The dimension of Ω_{j^*+2} is at least the dimension of the span of the covectors: $\mathcal{D}h,\ \mathcal{Z}'_0,\ \mathcal{Z}'_1,\ \cdots,\ \mathcal{Z}'_{j^*-1}$. From the definition of j^*, we know that the vectors $\mathcal{D}h,\ \mathcal{D}\mathcal{B}_0,\ \mathcal{D}\mathcal{B}_1,\ \cdots,\ \mathcal{D}\mathcal{B}_{j^*-1}$ are independent meaning that the dimension of their span is j^*+1. Hence, from (6.21), it easily follows that the dimension of the span of the vectors $\mathcal{D}h,\ \mathcal{Z}'_0,\ \mathcal{Z}'_1,\ \cdots,\ \mathcal{Z}'_{j^*-1},\ \mathcal{D}\tau$ is at least j^*+1. Since Ω_{j^*+3} contains this span, its dimension is at least j^*+1. Therefore, the condition $\Omega_{m+1} = \Omega_m$, for $m \geq j^*+3$ is achieved for $m \leq n+2$.

6.2.3 Extension to the case of multiple known inputs

It is immediate to repeat all the steps carried out in the previous two subsections and extend the validity of theorem 1 to the case of multiple known inputs ($m_u > 1$). Additionally, also theorem 2 can be easily extended to cope with the case of multiple known inputs. In this case, requiring that $\Omega_{m+1} = \Omega_m$ means that Ω_m must be invariant with respect to $\mathcal{L}_{\frac{g}{L_g^1}}$ and all $\mathcal{L}_{f^i}$ simultaneously.

6.3 Proof of the validity of the proposed analytic method in the general case

This section proves the analytic results presented in chapters 3 and 4 in the general case, i.e., $\forall m_w$ and when the dynamics are affine in the inputs (and not simply linear). In other words, for general nonlinear systems characterized by (2.1).

As in the previous section, we start by considering the case of a single known input ($m_u = 1$) and then we extend the obtained results to the case of several known inputs ($m_u > 1$).

In this setting, the observable codistribution for the extended system $\Sigma^{(k)}$ is obtained by algorithm 7, where the vector field G is the one in (6.3) and the symbol $\mathcal{D}_E$ denotes the differential with respect to the state in (6.1).

We remind the reader that, to reduce notational complexity, we adopt the Einstein notation. According to this notation, a dummy index is summed over. When the dummy index is Latin, the sum is from 1 to m_w. When it is Greek, the sum is from 0 to m_w. In addition, to refer to the components of a tensor, e.g., Γ^α_k, $\alpha = 0, 1, \cdots, m_w$ and $k = 1, \cdots, m_w$, we simply write Γ^α_k, $\forall \alpha, k$.

6.3.1 Observable Codistribution

In the general case, we do not have the same result of separation stated by theorem 1. This fact, however, does not prevent us to derive a codistribution that only depends on the original state and that fully characterizes the observability properties. This is the codistribution generated by algorithm 5, which is convergent, as it is proven in section 6.3.2.

In appendix A we prove that any system that satisfies (2.1) can be either set in canonic form, or some of the unknown inputs are spurious, i.e., they can be eliminated in order to derive the system observability properties. In the latter case, the system can be set in canonic form with respect to the remaining unknown inputs. Therefore, without loss of generality, we assume that our system is in canonic form. We select the scalar functions $h_1, \cdots, h_{m_w}$ such that their differentials ($\mathcal{D}_w$, i.e., the differential with respect to the unknown inputs vector) span the codistribution $\mathcal{D}_w \mathcal{L}_G \mathcal{F}$, which is defined in appendix A. Note that, as mentioned at the beginning of appendix A, the codistribution $span\{\mathcal{D}_E h_1, \ \cdots, \ \mathcal{D}_E h_{m_w}\} \subseteq \bar{\Omega}_j$, for a given integer j.

For each integer m, we generate the codistribution Ω_m by using algorithm 5 (note that here we are considering $m_u = 1$). By construction, the generators of Ω_m are the differentials of scalar functions that only depend on the original state (x) and not on its extension. In the sequel, we need to embed this codistribution in $\mathbb{R}^{n+m_w k}$. We will denote by $[\Omega_m, 0_{km_w}]$ the codistribution made by covectors whose first n components are covectors in Ω_m and the last components are all zero.

Additionally, we will denote by $L^m_{m_w}$ the codistribution that is the span of the Lie derivatives of $\mathcal{D}_E h_1, \ \cdots, \ \mathcal{D}_E h_{m_w}$ up to the order m along the vector G, i.e., $L^m_{m_w} \triangleq$

$$span\left\{\mathcal{L}^1_G \mathcal{D}_E h_1, \cdots, \mathcal{L}^1_G \mathcal{D}_E h_{m_w}, \cdots, \mathcal{L}^m_G \mathcal{D}_E h_1, \cdots, \mathcal{L}^m_G \mathcal{D}_E h_{m_w}\right\}$$

We finally introduce the following codistribution:

Definition 6 ($\tilde{\Omega}$ codistribution) *This codistribution is defined as follows:* $\tilde{\Omega}_m \triangleq [\Omega_m, 0_{km_w}] + L^m_{m_w}$

The codistribution $\tilde{\Omega}_m$ consists of two parts. Specifically, we can select a basis that consists of exact differentials that are the differentials of functions that only depend on the original state (x) and not on its extension (these are the generators of $[\Omega_m, 0_{m_w k}]$) and the differentials

$\mathcal{L}^1_G\mathcal{D}_Eh_1,\cdots,\mathcal{L}^1_G\mathcal{D}_Eh_{m_w},\cdots,\mathcal{L}^m_G\mathcal{D}_Eh_1,\cdots,\mathcal{L}^m_G\mathcal{D}_Eh_{m_w}$. The second set of generators, i.e., the differentials $\mathcal{L}^1_G\mathcal{D}_Eh_1,\cdots,\mathcal{L}^1_G\mathcal{D}_Eh_{m_w},\cdots,\mathcal{L}^m_G\mathcal{D}_Eh_1,\cdots,\mathcal{L}^m_G\mathcal{D}_Eh_{m_w}$, are mm_w and, with respect to the first set, they are differentials of functions that also depend on the state extension $\xi = [w,\ w^{(1)},\cdots,\ w^{(m-1)}]^T$. We have the following result, which is exactly the extension of lemma 2:

Lemma 11 *Let us denote with x_j the j^{th} component of the state ($j = 1,\cdots,n$). We have: $\mathcal{D}x_j \in \Omega_m$ if and only if $\mathcal{D}_Ex_j \in \tilde{\Omega}_m$*

Proof: As for lemma 2, the fact that $\mathcal{D}x_j \in \Omega_m$ implies that $\mathcal{D}_Ex_j \in \tilde{\Omega}_m$ is obvious since $[\Omega_m, 0_{m_wk}] \subseteq \tilde{\Omega}_m$ by definition. Let us prove that also the contrary holds, i.e., that if $\mathcal{D}_Ex_j \in \tilde{\Omega}_m$ then $\mathcal{D}x_j \in \Omega_m$. Since $\mathcal{D}_Ex_j \in \tilde{\Omega}_m$ we have $\mathcal{D}_Ex_j = \sum_{i=1}^{N_1} c^1_i\omega^1_i + \sum_{i=1}^{N_2} c^2_i\omega^2_i$, where $\omega^1_1,\omega^1_2,\cdots,\omega^1_{N_1}$ are N_1 generators of $[\Omega_m, 0_{m_wk}]$ and $\omega^2_1,\omega^2_2,\cdots,\omega^2_{N_2}$ are N_2 generators of $L^m_{m_w}$. We want to prove that $N_2 = 0$.

We proceed by contradiction. Let us suppose that $N_2 \geq 1$. We remark that the first set of generators have the last m_wk entries equal to zero, as for $\mathcal{D}_Ex_j$. The second set of generators consists of the Lie derivatives of $\mathcal{D}_Eh_1$ and $\mathcal{D}_Eh_{m_w}$ along G up to the m order. Let us select the following m_w generators among this second set. They are the highest order Lie derivative along G of $\mathcal{D}_Eh_1$ and $\mathcal{D}_Eh_{m_w}$, respectively. Let us denote by $j'_1,\cdots,j'_{m_w}$ these highest orders and let us denote by j' the largest value among $j'_1,\cdots,j'_{m_w}$. This means that at least one of the integers among $j'_1,\cdots,j'_{m_w}$ is equal to j'. We denote by $j'_{m_1},\cdots,j'_{m_r}$ all the $r(\geq 1)$ highest orders that are equal to j'. It is immediate to realize that $\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_1},\cdots,\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_r}$ are the only generators among the N_2 generators of above whose entries between the $(n+(j'-1)m_w+1)^{th}$ and $(n+j'm_w)^{th}$ can be different from zero. This because the functions $\mathcal{L}^{j'}_Gh_{m_1},\cdots,\mathcal{L}^{j'}_Gh_{m_r}$ are the only ones that depend on $w^{(j'-1)}$. By a direct computation, we can derive this dependency. We easily obtain that $\mathcal{L}^{j'}_Gh_{m_j}$ ($j = 1,\cdots,r$) depends on $w^{(j'-1)}$, by the following linear expression: $\mu^l_{m_j}w^{(j'-1)}_l$. Let us consider any linear combination of $\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_1},\cdots,\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_r}$, i.e., $\alpha^1\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_1}+\cdots+\alpha^r\mathcal{L}^{j'}_G\mathcal{D}_Eh_{m_r}$, with α non null. We remark that the function $\alpha^1\mathcal{L}^{j'}_Gh_{m_1}+\cdots+\alpha^r\mathcal{L}^{j'}_Gh_{m_r}$ depends on $w^{(j'-1)}$ as follows: $\sum_{j=1}^r \alpha^j\mu^l_{m_j}w^{(j'-1)}_l$. We also remark that it must exists at least one value of l such that $\sum_{j=1}^r \alpha^j\mu^l_{m_j} \neq 0$ (if this is not true it means that the tensor μ is singular). Hence, we obtain that $\mathcal{D}_Ex_j$ has at least one entry, among the last m_wk entries, different from zero and this is not possible ■

In the general case, we do not have the same result stated by theorem 1. We prove that the codistribution generated by algorithm 5, which is convergent (see section 6.3.1), fully characterizes the observability properties of the original state. This is proven by using the following three results, which holds for a scalar function $\theta(x)$ of the original state:

1. If $\theta(x)$ is weakly observable, $\exists m$ such that $\mathcal{D}_E\theta(x) \in \bar{\Omega}_m$;
2. $\bar{\Omega}_m \subseteq \tilde{\Omega}_m \triangleq [\Omega_m, 0_{km_w}] + L^m_{m_w}$;
3. If for a given integer m, $\mathcal{D}\theta(x) \in \Omega_m$ then $\theta(x)$ is weakly observable.

The first result is proven by proposition 6. In the rest of this section we prove the last two results (propositions 11 and 12).

We start by proving the second result (i.e., the one stated by proposition 11). This proof follows several steps, which are similar to the ones operated to prove theorem 1.

As in that case, for a given $m \leq k$ we define the vector $\Phi_m \in \mathbb{R}^{n+km_w}$ by the following algorithm:

1. $\Phi_0 = F$;

2. $\Phi_m = [\Phi_{m-1},\ G]$

where now the field G is the one given in (6.3) and the Lie brackets $[\cdot, \cdot]$ are computed with respect to the extended state, whose dimension is now $n + km_w$.

As in the case $m_w = 1,\ g^0 = 0$, it is immediate to realize that Φ_m has the last components (in this case km_w) identically null. In the sequel, we will denote by $\breve{\Phi}_m$ the vector in $\mathbb{R}^n$ that contains the first n components of Φ_m. In other words, $\Phi_m \triangleq [\breve{\Phi}_m^T, 0_{km_w}^T]^T$. Additionally, we set $\hat{\phi}_m \triangleq \begin{bmatrix} \phi_m \\ 0_{km_w} \end{bmatrix}$, where, for the brevity sake, here we denote by ϕ_m the vector $\phi_m^{\alpha_1,\cdots,\alpha_m}$ defined by algorithm 4 (we adopt this simplified notation when there is no ambiguity).

From now on, we adopt the Einstein notation in order to achieve notational brevity. According to this notation, dummy indexes from Latin alphabet implicate a sum from 1 to m_w (not explicitly written). Similarly, dummy indexes from Greek alphabet implicate a sum from 0 to m_w.

The result stated by lemma 3 still holds and the proof is identical. Also the result stated by lemma 4 still holds. However, the proof is more complicated. We have:

Lemma 12 *$\breve{\Phi}_m = \sum_{j=1}^m c^j_{\alpha_1,\cdots,\alpha_j}(\mathcal{L}_G h_1, \cdots, \mathcal{L}_G h_{m_w}, \cdots, \mathcal{L}_G^m h_1, \cdots, \mathcal{L}_G^m h_{m_w})\phi_j^{\alpha_1,\cdots,\alpha_j}$, i.e., the vector $\breve{\Phi}_m$ is a linear combination of the vectors $\phi_j^{\alpha_1,\cdots,\alpha_j}$ $(j = 1, \cdots, m)$, where the coefficients $(c^j_{\alpha_1,\cdots,\alpha_j})$ depend on the state only through the functions that generate the codistribution $L^m_{m_w}$.*

Proof: We proceed by induction. By definition, $\breve{\Phi}_0 = \phi_0$.
Inductive step: Let us assume that

$$\breve{\Phi}_{m-1} = \sum_{j=1}^{m-1} c^j_{\alpha_1,\cdots,\alpha_j}(\mathcal{L}_G h_1, \cdots, \mathcal{L}_G h_{m_w}, \cdots, \mathcal{L}_G^m h_1, \cdots, \mathcal{L}_G^m h_{m_w})\phi_j^{\alpha_1,\cdots,\alpha_j}$$

We have:

$$\Phi_m = [\Phi_{m-1},\ G] = \sum_{j=1}^{m-1} \left[c^j_{\alpha_1,\cdots,\alpha_j} \begin{bmatrix} \phi_j^{\alpha_1,\cdots,\alpha_j} \\ 0_{km_w} \end{bmatrix},\ G \right] =$$

$$\sum_{j=1}^{m-1} c^j_{\alpha_1,\cdots,\alpha_j} \left[\begin{bmatrix} \phi_j^{\alpha_1,\cdots,\alpha_j} \\ 0_{km_w} \end{bmatrix},\ G \right] - \mathcal{L}_G c^j_{\alpha_1,\cdots,\alpha_j} \begin{bmatrix} \phi_j^{\alpha_1,\cdots,\alpha_j} \\ 0_{km_w} \end{bmatrix} \tag{6.25}$$

We directly compute the Lie bracket in the sum (note that $\phi_j^{\alpha_1,\cdots,\alpha_j}$ is independent of the unknown input vector w and its time derivatives):

$$\left[\begin{bmatrix} \phi_j^{\alpha_1,\cdots,\alpha_j} \\ 0_{km_w} \end{bmatrix},\ G \right] = \begin{bmatrix} [\phi_j^{\alpha_1,\cdots,\alpha_j},\ g^0] + [\phi_j^{\alpha_1,\cdots,\alpha_j},\ g^i] w_i \\ 0_{km_w} \end{bmatrix} \tag{6.26}$$

On the other hand, by using (3.3) (3.4) and (3.6) it is immediate to obtain:

$$w_i = \nu_i^k \mathcal{L}_G h_k - \nu_i^k \mu_k^0, \qquad \forall i \tag{6.27}$$

and by substituting (6.27) in (6.26) and by using (3.10) we obtain

$$[\phi_j^{\alpha_1,\cdots,\alpha_j},\ g^0] + [\phi_j^{\alpha_1,\cdots,\alpha_j},\ g^i] w_i = [\phi_j^{\alpha_1,\cdots,\alpha_j}]^0 + [\phi_j^{\alpha_1,\cdots,\alpha_j}]^k \mathcal{L}_G h_k =$$

$$\phi_{j+1}^{\alpha_1,\cdots,\alpha_j,0} + \phi_{j+1}^{\alpha_1,\cdots,\alpha_j,k}\mathcal{L}_G h_k, \quad \forall \alpha_1,\cdots,\alpha_j \tag{6.28}$$

The last equality is due to the definition of the fields $\phi_{j+1}^{\alpha_1,\cdots,\alpha_j,\alpha_{j+1}}$, provided by algorithm 4.

Regarding the second term in (6.25), we remark that $\mathcal{L}_G c^j_{\alpha_1,\cdots,\alpha_j} = \sum_{k=1}^{m-1} \frac{\partial c^j_{\alpha_1,\cdots,\alpha_j}}{\partial(\mathcal{L}^k_G h_i)} \mathcal{L}_G^{k+1} h_i$. By substituting this equality and the one in (6.28) in (6.25) and by proceeding exactly as in the last part of the proof of lemma 4, we easily obtain the proof of the statement ■
In the general case, the result stated by lemma 5 does not hold, and this is one of the reasons because the separation property (theorem 1) does not hold. In particular, by only using lemma 12 and not the analogous of lemma 5, we cannot prove proposition 7. On the other hand, by only using lemma 12 it is immediate to obtain the following weaker result:

Proposition 9 *For any scalar function $h(x)$ we have:*

$$span\{\mathcal{L}_{\Phi_0}\mathcal{D}_E h, \mathcal{L}_{\Phi_1}\mathcal{D}_E h, \cdots, \mathcal{L}_{\Phi_m}\mathcal{D}_E h\} + L^m_{m_w} \subseteq span\{\mathcal{L}_{\hat{\phi}_0}\mathcal{D}_E h, \mathcal{L}_{\hat{\phi}_1^{\alpha_1}}\mathcal{D}_E h, \cdots, \mathcal{L}_{\hat{\phi}_m^{\alpha_1,\cdots,\alpha_m}}\mathcal{D}_E h\} + L^m_{m_w}$$

We also have the following result (note that also this result becomes an equality when $m_w = 1,\ g^0 \neq 0,\ w \neq 0$ and has been used in the proof of theorem 1):

Proposition 10

$$L^1_{m_w} + \mathcal{L}_G[\Omega_{m-1}, 0_{km_w}] \subseteq L^1_{m_w} + [\mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1}, 0_{km_w}]$$

Proof: Let us consider a generic covector in Ω_{m-1}. By construction, we know that it exists a scalar function $\theta(x)$ such that this covector is proportional to $\mathcal{D}\theta$. We have:

$$\mathcal{L}_G\theta = \mathcal{L}_{g^0}\theta + \mathcal{L}_{g^i}\theta w_i$$

and, by using (6.27), we have:

$$\mathcal{L}_G\theta = \mathcal{L}_{g^0}\theta + \mathcal{L}_{g^i}\theta(\nu_i^k\mathcal{L}_G h_k - \nu_i^k\mu_k^0)$$

and, by using (3.9), we obtain:

$$\mathcal{L}_G\theta = \mathcal{L}_{\hat{g}^0}\theta + \mathcal{L}_{\hat{g}^k}\theta\mathcal{L}_G h_k$$

from which the proof follows ■
We are now ready to prove proposition 11.

Proposition 11 *For any integer m, $\bar{\Omega}_m \subseteq \tilde{\Omega}_m \triangleq [\Omega_m, 0_{km_w}] + L^m_{m_w}$.*

Proof: This proof follows the same steps of the proof of theorem 1. We proceed by induction. By definition, $\bar{\Omega}_0 = \tilde{\Omega}_0$ since they are both the span of the output (or the outputs in case of multiple outputs).
Inductive step: Let us assume that $\bar{\Omega}_{m-1} \subseteq \tilde{\Omega}_{m-1}$. We have: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_F\bar{\Omega}_{m-1} + \mathcal{L}_G\bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + \mathcal{L}_F\tilde{\Omega}_{m-1} + \mathcal{L}_G\bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_F L^{m-1}_{m_w} + \mathcal{L}_G\bar{\Omega}_{m-1}$.
On the other hand,

$$\mathcal{L}_F L^{m-1}_{m_w} = \mathcal{L}_F\mathcal{L}^1_G\mathcal{D}_E h_i + \cdots + \mathcal{L}_F\mathcal{L}_G^{m-2}\mathcal{D}_E h_i + \mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h_i$$

The only terms which are not in $\bar{\Omega}_{m-1}$ are $\mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h_i$.
Hence we have:

$$\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_F\mathcal{L}_G^{m-1}\mathcal{D}_E h_i + \mathcal{L}_G\bar{\Omega}_{m-1}$$

By using lemma 3 we obtain:

$$\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h_i + \mathcal{L}_G\bar{\Omega}_{m-1}$$

By using again the induction assumption we obtain:

$$\bar{\Omega}_m = [\Omega_{m-1}, 0_{km_w}] + L_{m_w}^{m-1} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h_i + \mathcal{L}_G[\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_G L_{m_w}^{m-1} =$$

$$[\Omega_{m-1}, 0_{km_w}] + L_{m_w}^{m} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_{\Phi_{m-1}}\mathcal{D}_E h_i + \mathcal{L}_G[\Omega_{m-1}, 0_{km_w}]$$

and by using propositions 9 and 10 we obtain: $\bar{\Omega}_m \subseteq [\Omega_{m-1}, 0_{km_w}] + L_{m_w}^{m} + [\mathcal{L}_f\Omega_{m-1}, 0_{km_w}] + \mathcal{L}_{\hat{\phi}_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\mathcal{D}_E h_i + [\mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1}, 0_{km_w}] = \tilde{\Omega}_m$ ■

Before proceeding with the proof of proposition 12, we need to prove the following lemma:

Lemma 13 *Let us consider the scalar function $\lambda(x,w) = \theta(x) + \theta^i(x)\mathcal{L}_G h_i$ and let us assume that it is observable in Σ^1. Then, all the m_w+1 functions $\theta(x), \theta^1(x), \cdots, \theta^{m_w}(x)$ are observable in $\Sigma^{(0)}$.*

Proof: Let us consider two points x_a and x_b, indistinguishable in $\Sigma^{(0)}$. We have to prove that $\theta(x_a) = \theta(x_b)$ and $\theta^i(x_a) = \theta^i(x_b)$ $(\forall i)$. From definition 3 and the definition of a non-shy set, we know that there exist m_w+1 distinct pairs of vectors (w_a^0, w_b^0), (w_a^1, w_b^1), $\cdots$, $(w_a^{m_w}, w_b^{m_w})$, such that the two points $[x_a, w_a^\alpha]$ and $[x_b, w_b^\alpha]$, $\forall\alpha$, are indistinguishable in $\Sigma^{(1)}$, and the vector space generated by the vectors $t^j \triangleq w_a^j - w_a^0$ $(\forall j)$ has dimension m_w.

From the fact that the function λ is observable and by using the previous m_w+1 indistinguishable points we obtain the following m_w+1 equations:

$$\lambda(x_a, w_a^\alpha) = \lambda(x_b, w_b^\alpha), \quad \forall\alpha$$

By using the expression $\lambda(x,w) = \theta(x) + \theta^i(x)\mathcal{L}_G h_i$, the fact that also the functions $\mathcal{L}_G h_i$ are observable we obtain:

$$\theta(x_a) + \theta^i(x_a)\mathcal{L}_G h_i(x_a, w_a^\alpha) = \theta(x_b) + \theta^i(x_b)\mathcal{L}_G h_i(x_a, w_a^\alpha), \quad \forall\alpha$$

By using the expression $\mathcal{L}_G h_i = \mu_i^0 + \mu_i^k w_k$ we obtain:

$$\theta(x_a) + \theta^i(x_a)[\mu_i^0(x_a) + \mu_i^k(x_a)(w_a^\alpha)_k] = \theta(x_b) + \theta^i(x_b)[\mu_i^0(x_a) + \mu_i^k(x_a)(w_a^\alpha)_k], \quad \forall\alpha$$

By subtracting from the last m_w equations the first equation we obtain:

$$\theta^i(x_a)\mu_i^k(x_a)(w_a^j - w_a^0)_k = \theta^i(x_b)\mu_i^k(x_a)(w_a^j - w_a^0)_k, \quad \forall j$$

namely,

$$\mu_i^k t_k^j(\theta^i(x_a) - \theta^i(x_b)) = 0, \quad \forall j$$

Since both μ_i^k and t_k^j are non singular, $\theta^i(x_a) = \theta^i(x_b)$. By using the first equation $(\alpha = 0)$ we also obtain $\theta(x_a) = \theta(x_b)$ ■

We are now ready to prove the following result:

Proposition 12 *For any integer m, the codistribution $\tilde{\Omega}_m$ is observable.*

Proof: We proceed by induction. By definition, $\tilde{\Omega}_0$ is the span of the differentials of the observable functions $h_1, \cdots, h_{m_w}$ (and, in case of multiple outputs that do not appear in the selection, of these further outputs).
Inductive step: Let us assume that the codistribution $\tilde{\Omega}_{m-1}$ is observable. We want to prove that also $\tilde{\Omega}_m$ is observable. From algorithm 5, we need to prove that the following codistributions are observable:

$$\mathcal{L}_f\Omega_{m-1}, \quad \mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1}, \quad span\{\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\mathcal{D}_E h_i\}$$

$\forall\alpha, \alpha_1, \cdots, \alpha_{m-1}, i$.

Since $\tilde{\Omega}_{m-1}$ is observable, also $\mathcal{L}_F\tilde{\Omega}_{m-1}$ it is. Hence, $\mathcal{L}_F[\Omega_{m-1}, 0_{km_w}]$ is observable and, consequently, we obtain that $\mathcal{L}_f\Omega_{m-1}$ is observable.

We also have that $\mathcal{L}_G\tilde{\Omega}_{m-1}$ is observable and consequently, also $\mathcal{L}_G[\Omega_{m-1}, 0_{km_w}]$ it is. Let us consider a function $\theta(x)$ such that $\mathcal{D}_E\theta \in \Omega_{m-1}$. We obtain that $\mathcal{D}_E\mathcal{L}_G\theta$ belongs to the observable codistribution. On the other hand we have:

$$\mathcal{L}_G\theta = \mathcal{L}_{g^0}\theta + \mathcal{L}_{g^i}\theta w_i = \mathcal{L}_{g^0}\theta + \mathcal{L}_{g^i}\theta(\nu_i^k\mathcal{L}_G h_k - \nu_i^k\mu_k^0) = \mathcal{L}_{\hat{g}^0}\theta + \mathcal{L}_{\hat{g}^i}\theta\mathcal{L}_G h_i$$

where we used (6.27). From lemma 13 we immediately obtain that the codistribution $\mathcal{L}_{\hat{g}^\alpha}\Omega_{m-1}$ is observable $\forall\alpha$. It remains to show that also the codistribution $span\{\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\mathcal{D}_E h_i\}$ is observable $\forall\alpha, \alpha_1, \cdots, \alpha_{m-1}, i$.

To prove this, we start by remarking that, by applying $\mathcal{L}_F$ and $\mathcal{L}_G$ repetitively on the observable codistribution, starting from $[\Omega_{m-1},\ 0_{km_w}]$ and by proceeding as before, we finally obtain an observable codistribution in the space of the original state, which is invariant under $\mathcal{L}_{\hat{g}^\alpha}$ $(\forall\alpha)$ and $\mathcal{L}_f$. As in the case $m_w = 1$, $g^0 = 0$ (see proposition 8), it is possible to show that this codistribution is also invariant under $\mathcal{L}_{\phi_{m-2}^{\alpha_1,\cdots,\alpha_{m-2}}}$ $(\forall\alpha_1, \cdots, \alpha_{m-2})$. This means that the function $\mathcal{L}_{[\hat{\phi}_{m-2},\ G]}h_i$ is observable $\forall i$. Let us compute the Lie bracket of $\hat{\phi}_{m-2} = [\phi_{m-2}^{\alpha_1,\cdots,\alpha_{m-2}}, 0_{km_w}]^T$ with G (for the brevity sake we omit the first $m-2$ indexes (α)). We have:

$$[\hat{\phi}_{m-2},\ G] = \begin{bmatrix} [\phi_{m-2},\ g^0] \\ 0_{km_w} \end{bmatrix} + \begin{bmatrix} [\phi_{m-2},\ g^i]w_i \\ 0_{km_w} \end{bmatrix}$$

By using (6.27) we obtain:

$$[\hat{\phi}_{m-2},\ G] = \begin{bmatrix} [\phi_{m-2},\ g^0] + [\phi_{m-2},\ g^i](\nu_i^k\mathcal{L}_G h_k - \nu_i^k\mu_k^0) \\ 0_{km_w} \end{bmatrix} = \begin{bmatrix} [\phi_{m-2}]^0 + [\phi_{m-2}]^k\mathcal{L}_G h_k \\ 0_{km_w} \end{bmatrix} =$$

$$\begin{bmatrix} \phi_{m-1}^{\cdots,0} + \phi_{m-1}^{\cdots,k}\mathcal{L}_G h_k \\ 0_{km_w} \end{bmatrix}$$

Hence we have:

$$\mathcal{L}_{[\hat{\phi}_{m-2},\ G]}h_i = \mathcal{L}_{\phi_{m-1}^{\cdots,0}}h_i + \mathcal{L}_{\phi_{m-1}^{\cdots,k}}h_i\mathcal{L}_G h_k$$

and the observability of $\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\mathcal{D}_E h_i$, $\forall\alpha_1, \cdots, \alpha_{m-1}, i$, follows from lemma 13 ■

The results of this section can be summarized by the following theorem:

Theorem 3 (Observable Codistribution) *$\theta(x)$ is weakly observable iff $\exists m$ such that $\mathcal{D}\theta(x) \in \Omega_m$.*

Proof: If $\theta(x)$ is weakly observable, from proposition 6 we have that $\exists m$ such that $\mathcal{D}_E\theta(x) \in \bar{\Omega}_m$. From proposition 11 and lemma 11 we obtain that $\mathcal{D}\theta(x) \in \Omega_m$. Conversely, if for a given integer m, $\mathcal{D}\theta(x) \in \Omega_m$ then, from proposition 12 and lemma 11, $\theta(x)$ is weakly observable ■

6.3.2 Convergence

The goal of this section is to investigate the convergence properties of algorithm 5. We will show that the algorithm converges in a finite number of steps and we will also provide the criterion to establish that the algorithm has converged (theorem 4). This theorem will be proved at the end of this section since we need to introduce several important new quantities and properties.

When investigating the convergence properties of algorithm 5, we remark that, the main difference between algorithm 1 and 5, is the presence of the last term in the recursive step of the latter. Without this term, the convergence criterion would simply consist of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1.

The following result provides the convergence criterion in a very special case that basically occurs when the contribution due to the last term in the recursive step of algorithm 5 is included in the other terms. In this case, we obviously obtain that the convergence criterion consists of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1.
We have the following result:

Lemma 14 *Let us denote by Λ_j the distribution generated by $\phi_0, \phi_1^{\alpha_1}, \cdots, \phi_j^{\alpha_1,\cdots,\alpha_j}$, $\forall \alpha_1, \cdots, \alpha_j$ and let us denote by $m(\leq n-1)$ the smallest integer for which $\Lambda_{m+1} = \Lambda_m$ (n is the dimension of the state x). In the very special case when $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}} \mathcal{L}_{g^\alpha} h_k = 0$, $\forall j = 0, \cdots, m$, $\forall \alpha, \alpha_1, \cdots, \alpha_j, k$, algorithm 5 converges at the integer j such that $\Omega_{j+1} = \Omega_j$ and this occurs in at most $n-1$ steps.*

Proof: First of all, we remark that the existence of an integer $m(\leq n-1)$ such that $\Lambda_{m+1} = \Lambda_m$ is a simple extension of the result proved in [27], when the extended Lie bracket defined in definition 1 are adopted instead of the simple Lie bracket. In particular, it is possible to prove that the distribution converges to Λ^* and that the convergence is achieved at the smallest integer for which we have $\Lambda_{m+1} = \Lambda_m$. Additionally, we have $\Lambda_{m+1} = \Lambda_m = \Lambda^*$ and m cannot exceed $n-1$.

In the very special case when $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}} \mathcal{L}_{g^\alpha} h_k = 0$, $\forall j = 0, \cdots, m$, $\forall \alpha, \alpha_1, \cdots, \alpha_j, k$, thanks to the aforementioned convergence of the distribution Λ_j, we easily obtain that $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}} \mathcal{L}_{g^\alpha} h_k = 0$ $\forall j \geq 1$.

We have (by avoiding to write all the upper indexes) $\phi_j = [\phi_{j-1}]^{\alpha_j} = \nu_\beta^{\alpha_j} [\phi_{j-1},\ g^\beta]$. Hence:

$$\mathcal{L}_{\phi_j} h_m = \nu_\beta^{\alpha_j} \left(\mathcal{L}_{\phi_{j-1}} \mathcal{L}_{g^\beta} h_m - \mathcal{L}_{g^\beta} \mathcal{L}_{\phi_{j-1}} h_m \right) = \nu_\beta^{\alpha_j} \mathcal{L}_{\phi_{j-1}} \mathcal{L}_{g^\beta} h_m - \mathcal{L}_{\hat{g}^{\alpha_j}} \mathcal{L}_{\phi_{j-1}} h_m$$

Since $\mathcal{L}_{\phi_{j-1}} \mathcal{L}_{g^\beta} h_m = 0$ $\forall j \geq 1$, we have $\mathcal{L}_{\phi_j} h_m = -\mathcal{L}_{\hat{g}^{\alpha_j}} \mathcal{L}_{\phi_{j-1}} h_m$, for any $j \geq 1$. Therefore, in this case the last term in the recursive step of algorithm 5, is included in the second last term. Therefore, we conclude that algorithm 5 converges when $\Omega_{m+1} = \Omega_m$. This occurs in at most $n-1$ steps, as for algorithm 1. ■

Let us consider now the general case. To proceed we need to introduce several important new quantities and properties. As in the case $m_w = 1$, $g^0 = 0$ we introduce the analogous of the

field ψ. We start by defining the following new operation. Given a vector field a we define the following $m_w + 1$ vector fields as follows:

$$\{a\}^\alpha \triangleq [a,\ \hat{g}^\alpha] = [a]^\alpha + (\mathcal{L}_a \nu_j^\alpha)\mu_l^j \hat{g}^l \tag{6.29}$$

For a given integer $0 \leq m \leq k$ we define $(m_w+1)^m$ vectors $\phi_m^{\alpha_1,\cdots,\alpha_m} \in \mathbb{R}^n$ $(\forall \alpha_j)$ by the following recursive algorithm:

1. $\psi_0 = f$;
2. $\psi_m^{\alpha_1,\cdots,\alpha_m} = \{\psi_m^{\alpha_1,\cdots,\alpha_{m-1}}\}^{\alpha_m}$

It is possible to find a useful expression that relates these vectors to the vectors ϕ_j, previously defined. This is the analogous of the relation given in lemma 7. On the other hand, we provide here a simplified version of this relation. For the sequel, we do not need the complete relation. We have:

Lemma 15 *For $m \geq 1$, it holds the following equation:*

$$\psi_m^{\alpha_1,\cdots,\alpha_m} = \phi_m^{\alpha_1,\cdots,\alpha_m} + \sum_{k=1}^{k_m} \gamma_k \hat{j}^k \tag{6.30}$$

where:

- *k_m is a strictly positive integer that depends on m;*
- *the vectors $\hat{j}^k$ are among the vectors $\hat{g}^\alpha$ $(\forall\alpha)$ and their Lie brackets up to $m-1$ times;*
- *γ_k are scalar functions of the state x that satisfy $\mathcal{D}\gamma_k \in \Omega_{m+1}$.*

Before providing the proof of this lemma, we remark that equation (6.15), which holds in the special case $m_w = 1,\ g^0 = 0$, agrees with (6.30). In particular, in (6.15), the scalar functions γ_k are explicitly computed (and, because of lemma 8, their differentials belong to Ω_{m+1}). Regarding the vectors $\hat{j}^k$, in the case $m_w = 1,\ g^0 = 0$, we only have the vector $\hat{g}^1 \triangleq \frac{g^1}{L_g^1}$, since all the Lie brackets of this vector with itself vanish.

Proof: We proceed by induction. We first need to consider $m = 1$ and we compute the equation that relates ψ_1^α to ϕ_1^α, $\forall\alpha$. For $\alpha = i = 1, \cdots, m_w$ we have:

$$\psi_1^i = \{f\}^i = [f]^i + (\mathcal{L}_f \nu_j^i)\mu_l^j \hat{g}^l = \phi_1^i + (\mathcal{L}_f \nu_j^i)\mu_l^j \hat{g}^l$$

This equation agrees with (6.30), provided that we are able to prove that the differentials of all the components of the two-index tensor $(\mathcal{L}_f \nu_j^i)\mu_l^j$ belong to $\Omega_2 (= \Omega_{m+1})$. To prove this, we consider the operator $\mathcal{L}_{(\mathcal{L}_f \nu_j^i)\mu_l^j \hat{g}^l}$ From the previous equation, we easily obtain:

$$\mathcal{L}_{(\mathcal{L}_f \nu_j^i)\mu_l^j \hat{g}^l} = \mathcal{L}_{\psi_1^i} - \mathcal{L}_{\phi_1^i}$$

We apply this operator on h_k $(\forall k)$. The result is a function whose differential belong to $\Omega_2 (= \Omega_{m+1})$, by construction. We obtain:

$$\mathcal{L}_{(\mathcal{L}_f \nu_j^i)\mu_l^j \hat{g}^l} h_k = (\mathcal{L}_f \nu_j^i)\mu_l^j \mathcal{L}_{\hat{g}^l} h_k = (\mathcal{L}_f \nu_j^i)\mu_l^j \delta_k^l = (\mathcal{L}_f \nu_j^i)\mu_k^j$$

Hence, the differentials of all the components of the two-index tensor $(\mathcal{L}_f \nu^i_j)\mu^j_k$ belong to $\Omega_2(=\Omega_{m+1})$ and this proves the case $m=1,\ \alpha=i=1,\cdots,m_w$. Let us consider the case $m=1,\ \alpha=0$. We proceed in the same way. We have:

$$\psi^0_1=\{f\}^0=[f]^0-(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l=\phi^0_1-(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l$$

Again, we remark that this equation agrees with (6.30), provided that we are able to prove that the differentials of all the components of the one-index tensor $(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l$ belong to $\Omega_2(=\Omega_{m+1})$. To prove this, we consider the operator $\mathcal{L}_{(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l}$. From the previous equation, we easily obtain:

$$\mathcal{L}_{(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l}=\mathcal{L}_{\phi^0_1}-\mathcal{L}_{\psi^0_1}$$

We apply this operator on h_k $(\forall k)$. The result is a function whose differential belong to $\Omega_2(=\Omega_{m+1})$, by construction. We obtain:

$$\mathcal{L}_{(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l}h_k=(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\mathcal{L}_{\hat{g}^l}h_k=(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_l\delta^l_k=(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_k$$

Hence, the differentials of all the components of the one-index tensor $(\mathcal{L}_f\nu^i_j\mu^0_i)\mu^j_k$ belong to $\Omega_2(=\Omega_{m+1})$ and this proves the case $m=1,\ \alpha=0$.
Inductive step: Let us assume that (6.30) holds for $m-1$. We need to prove that:

$$\psi^{\alpha_1,\cdots,\alpha_{m-1},\alpha_m}_m=\phi^{\alpha_1,\cdots,\alpha_{m-1},\alpha_m}_m+\sum_{k=1}^{k_m}\gamma_k\hat{j}^k \tag{6.31}$$

with $\mathcal{D}\gamma_k\in\Omega_{m+1}$. Let us start by considering the case when $\alpha_m=i_m=1,\cdots,m_w$. We have:

$$\psi^{\alpha_1,\cdots,\alpha_{m-1},i_m}_m=\{\psi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}\}^{i_m}=\left\{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}+\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^{i_m}=$$

$$=\left\{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}\right\}^{i_m}+\left\{\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^{i_m} \tag{6.32}$$

Let us consider the second term:

$$\left\{\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^{i_m}=\left[\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k,\ \hat{g}^{i_m}\right]=$$

$$=\sum_{k=1}^{k_{m-1}}\gamma_k\left[\hat{j}^k,\ \hat{g}^{i_m}\right]-\sum_{k=1}^{k_{m-1}}(\mathcal{L}_{\hat{g}^{i_m}}\gamma_k)\hat{j}^k=\sum_{k=1}^{k_m}\gamma_k\hat{j}^k$$

with $\mathcal{D}\gamma_k\in\Omega_{m+1}$ and the vectors $\hat{j}^k$ are among the vectors $\hat{g}^\alpha$ $(\forall\alpha)$ and their Lie brackets up to $m-1$ times. We also have:

$$\left\{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}\right\}^{i_m}=\phi^{\alpha_1,\cdots,i_m}_m+(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^{i_m}_j)\mu^j_l\hat{g}^l$$

Substituting this in (6.32) we obtain:

$$\psi_m^{\alpha_1,\cdots,i_m} = \phi_m^{\alpha_1,\cdots,i_m} + (\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j\hat{g}^l + \sum_{k=1}^{k_m}\gamma_k\hat{j}^k$$

This equation agrees with (6.30), provided that we are able to prove that the differentials of all the components of the $(m+1)$-index tensor $(\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j$ belong to Ω_{m+1}. To prove this, we consider the following operator:

$$\mathcal{L}_{(\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j\hat{g}^l}$$

We apply this operator on h_k ($\forall k$). The result is a function whose differential belong to Ω_{m+1}, by construction. We obtain:

$$\mathcal{L}_{(\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j\hat{g}^l}h_k = (\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j\mathcal{L}_{\hat{g}^l}h_k = (\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j\delta_k^l =$$

$$(\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_k^j$$

Hence, the differentials of all the components of the $(m+1)$-index tensor $(\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^{i_m})\mu_l^j$ belong to Ω_{m+1} and this proves the inductive step when $\alpha_m = i_m = 1,\cdots,m_w$.

To conclude, it remains to prove the inductive step when $\alpha_m = 0$. We proceed in the same way. We have:

$$\psi_m^{\alpha_1,\cdots,\alpha_{m-1},0} = \{\psi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}\}^0 = \left\{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}} + \sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^0 =$$

$$= \left\{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}\right\}^0 + \left\{\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^0 \tag{6.33}$$

Let us consider the second term:

$$\left\{\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k\right\}^0 = \left[\sum_{k=1}^{k_{m-1}}\gamma_k\hat{j}^k,\ \hat{g}^0\right] =$$

$$= \sum_{k=1}^{k_{m-1}}\gamma_k\left[\hat{j}^k,\ \hat{g}^0\right] - \sum_{k=1}^{k_{m-1}}(\mathcal{L}_{\hat{g}^0}\gamma_k)\hat{j}^k = \sum_{k=1}^{k_m}\gamma_k\hat{j}^k$$

with $\mathcal{D}\gamma_k \in \Omega_{m+1}$ and the vectors $\hat{j}^k$ are among the vectors $\hat{g}^\alpha$ ($\forall\alpha$) and their Lie brackets up to $m-1$ times. We also have:

$$\left\{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}\right\}^0 = \phi_m^{\alpha_1,\cdots,0} - (\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^i\mu_i^0)\mu_l^j\hat{g}^l$$

Substituting this in (6.33) we obtain:

$$\psi_m^{\alpha_1,\cdots,0} = \phi_m^{\alpha_1,\cdots,0} - (\mathcal{L}_{\phi_{m-1}^{\alpha_1,\cdots,\alpha_{m-1}}}\nu_j^i\mu_i^0)\mu_l^j\hat{g}^l + \sum_{k=1}^{k_m}\gamma_k\hat{j}^k$$

This equation agrees with (6.30), provided that we are able to prove that the differentials of all the components of the m-index tensor $(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l$ belong to Ω_{m+1}. To prove this, we consider the following operator:

$$\mathcal{L}_{(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l}$$

We apply this operator on h_k ($\forall k$). The result is a function whose differential belong to Ω_{m+1}, by construction. We obtain:

$$\mathcal{L}_{(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l\hat{g}^l}h_k = (\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l\mathcal{L}_{\hat{g}^l}h_k = (\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l\delta^l_k =$$

$$(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_k$$

Hence, the differentials of all the components of the m-index tensor $(\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^i_j\mu^0_i)\mu^j_l$ belong to Ω_{m+1} and this proves the inductive step also when $\alpha_m = 0$ ■

For any integer $m \geq 1$, we introduce two fundamental $(m+1)$-index tensors. They are both characterized by m upper indexes and 1 lower index. They are:

$$\mathcal{O}^{\alpha_1,\cdots,\alpha_{m-1},\alpha_m}_\gamma \triangleq (\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\nu^{\alpha_m}_\beta)\mu^\beta_\gamma$$

$$\mathcal{P}^{\alpha_1,\cdots,\alpha_{m-1},\alpha_m}_\gamma \triangleq (\mathcal{L}_{\phi^{\alpha_1,\cdots,\alpha_{m-1}}_{m-1}}\mu^\beta_\gamma)\nu^{\alpha_m}_\beta$$

For the brevity sake, we avoid to write all the indexes $\alpha_1,\cdots,\alpha_{m-1}$. The previous tensors can be written as follows:

$$\mathcal{O}^{\cdots,\alpha_m}_\gamma \triangleq (\mathcal{L}_{\phi_{m-1}}\nu^{\alpha_m}_\beta)\mu^\beta_\gamma \tag{6.34}$$

$$\mathcal{P}^{\cdots,\alpha_m}_\gamma \triangleq (\mathcal{L}_{\phi_{m-1}}\mu^\beta_\gamma)\nu^{\alpha_m}_\beta \tag{6.35}$$

We remark that $(\mathcal{L}_{\phi_{m-1}}\nu^\alpha_\beta\mu^\beta_\gamma) = (\mathcal{L}_{\phi_{m-1}}\delta^\alpha_\gamma) = 0$. Hence:

$$\mathcal{O}^{\cdots,\alpha}_\gamma = -\mathcal{P}^{\cdots,\alpha}_\gamma \tag{6.36}$$

The following result extends the one given in lemma 8:

Lemma 16 *For any integer $m \geq 1$, the differentials of all the components of the two $(m+1)$-index tensors given in (6.34) and (6.35) belong to Ω_{m+1}:*

Proof: By using equation (6.36), we only need to consider one of these tensors. Let us refer to $\mathcal{O}$. Let us start by considering the components with $\alpha_m = i_m = 1,\cdots,m_w$. The proof that the differentials of these components belong to Ω_{m+1} is available in the proof of lemma 15. Let us consider the components $\mathcal{O}^{\cdots,0}_k$. At the end of the proof of lemma 15, we proved that the differentials of $(\mathcal{L}_{\phi_{m-1}}\nu^i_j\mu^0_i)\mu^j_k$ belong to Ω_{m+1}. We have:

$$(\mathcal{L}_{\phi_{m-1}}\nu^i_j\mu^0_i)\mu^j_k = (\mathcal{L}_{\phi_{m-1}}\mu^0_i)\nu^i_j\mu^j_k + (\mathcal{L}_{\phi_{m-1}}\nu^i_j)\mu^j_k\mu^0_i = (\mathcal{L}_{\phi_{m-1}}\mu^0_i)\delta^i_k + \mathcal{O}^{\cdots,i}_k\mu^0_i = -\mathcal{O}^{\cdots,0}_k$$

■

The result stated by proposition 8 also holds in the general case. We have:

Proposition 13 *If Ω_m is invariant with respect to $\mathcal{L}_f$ and $\mathcal{L}_{\hat{g}^\alpha}$, $\forall\alpha$, then it is also invariant with respect to $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}}$ for $j=1,\cdots,m-1$.*

Proof: From (6.30) we obtain the following operator equality:

$$\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}} = \mathcal{L}_{\psi_j^{\alpha_1,\cdots,\alpha_j}} - \sum_{k=1}^{k_j} \mathcal{L}_{\gamma_k \hat{j}^k}$$

with $\mathcal{D}\gamma_k \in \Omega_{j+1}$. Let us apply the previous equality on a given covector in Ω_m. Because of the invariance with respect to $\mathcal{L}_f$ and $\mathcal{L}_{\hat{g}^\alpha}$, $\forall\alpha$, we also have the invariance with respect to $\mathcal{L}_{\hat{j}^k}$ and $\mathcal{L}_{\psi_j^{\alpha_1,\cdots,\alpha_j}}$. If $j \leq m-1$, $\mathcal{D}\gamma_k \in \Omega_m$ and we obtain the invariance with respect to $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}}$ $\blacksquare$
The following result extends the one given in lemma 9. As for the equation (6.17), we need to substitute the expression of ϕ_j in terms of ϕ_{j-2} in the term $\mathcal{L}_{\phi_j} h$. This will allow us to detect the key quantity that governs the convergence of algorithm 5, in particular regarding the contribution due to the last term in the recursive step of algorithm 5. In the case $m_w = 1$, $g^0 = 0$, this quantity was a scalar and it was the one provided in (3.2). In the general case, the derivation was very troublesome since we did not know a priori that, instead of a scalar, the key quantity becomes a three-index tensor. Specifically, it is the tensor defined by (3.11), of type $(2,1)$. For the sake of clarity, we provide equation (3.11) below:

$$\mathcal{T}_\gamma^{\alpha,\beta} \triangleq \nu_\eta^\beta(\mathcal{L}_{\hat{g}^\alpha}\mu_\gamma^\eta), \quad \forall\alpha,\beta,\gamma$$

In general, this tensor has $(m_w+1)\times(m_w+1)\times(m_w+1)$ components. On the other hand, in our coordinate setting (for which equations (3.7) and (3.8) hold) it is immediate to obtain that $\mathcal{T}_0^{\alpha,\beta} = 0$ $\forall\alpha,\beta$. In other words, in this setting we can consider the lower index as a Latin index and the components of this tensor are $(m_w+1)\times(m_w+1)\times m_w$.

We have the following fundamental analytic result:

Lemma 17 *For $j \geq 2$, we have the following key equality (for the brevity sake, we denote by three dots $(\cdots)$, the first $j-2$ Greek indexes, $\alpha_1,\cdots,\alpha_{j-2}$):*

$$\mathcal{L}_{\phi_j^{\cdots,\alpha_{j-1},\alpha_j}} h_m = \tag{6.37}$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j} - \mathcal{O}_k^{\cdots,\alpha_j}\mathcal{T}_m^{\alpha_{j-1},k} - \mathcal{O}_k^{\cdots,\alpha_{j-1}}\mathcal{T}_m^{k,\alpha_j} - \mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\cdots,\alpha_j} - \mathcal{T}_k^{\alpha_{j-1},\alpha_j}\mathcal{P}_m^{\cdots,k} - \mathcal{L}_{\hat{g}^{\alpha_j}}\mathcal{L}_{\phi_{j-1}} h_m$$

Proof: We will prove this equality by an explicit computation.

First case: We have $\phi_j = [\phi_{j-1}]^{\alpha_j} = \nu_\beta^{\alpha_j}[\phi_{j-1},\ g^\beta]$. Hence:

$$\mathcal{L}_{\phi_j} h_m = \nu_\beta^{\alpha_j}\left(\mathcal{L}_{\phi_{j-1}}\mathcal{L}_{g^\beta} h_m - \mathcal{L}_{g^\beta}\mathcal{L}_{\phi_{j-1}} h_m\right) = \nu_\beta^{\alpha_j}\mathcal{L}_{\phi_{j-1}}\mathcal{L}_{g^\beta} h_m - \mathcal{L}_{\hat{g}^{\alpha_j}}\mathcal{L}_{\phi_{j-1}} h_m \tag{6.38}$$

The second term coincides with the last term in (6.37). Hence, we need to prove that:

$$\nu_\beta^{\alpha_j}\mathcal{L}_{\phi_{j-1}}\mathcal{L}_{g^\beta} h_m = \mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j} - \mathcal{O}_k^{\cdots,\alpha_j}\mathcal{T}_m^{\alpha_{j-1},k} - \mathcal{O}_k^{\cdots,\alpha_{j-1}}\mathcal{T}_m^{k,\alpha_j} - \mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\cdots,\alpha_j} - \mathcal{T}_k^{\alpha_{j-1},\alpha_j}\mathcal{P}_m^{\cdots,k} \tag{6.39}$$

We have:

$$\nu_\beta^{\alpha_j}\mathcal{L}_{\phi_{j-1}}\mathcal{L}_{g^\beta} h_m = \nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\eta}\mathcal{L}_{g^\beta} h_m - \nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{g^\eta}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\beta} h_m \tag{6.40}$$

Let us compute these two terms separately. For the first we obtain:

$$\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\eta}\mathcal{L}_{g^\beta}h_m=\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\eta}\mu_m^\beta=\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{g^\eta}\mu_m^\beta)-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}})\mathcal{L}_{g^\eta}\mu_m^\beta=$$

$$\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mu_m^\beta)-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\nu_{\eta'}^{\alpha_{j-1}})\delta_\eta^{\eta'}\mathcal{L}_{g^\eta}\mu_m^\beta=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\nu_{\eta'}^{\alpha_{j-1}})\mu_\gamma^{\eta'}\nu_\eta^\gamma\mathcal{L}_{g^\eta}\mu_m^\beta=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j}\nu_{\eta'}^{\alpha_{j-1}})\mu_\gamma^{\eta'}\mathcal{L}_{\hat{g}^\gamma}\mu_m^\beta=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j})\nu_{\eta'}^{\alpha_{j-1}}\mu_\gamma^{\eta'}\mathcal{L}_{\hat{g}^\gamma}\mu_m^\beta-\mathcal{L}_{\phi_{j-2}}(\nu_{\eta'}^{\alpha_{j-1}})\nu_\beta^{\alpha_j}\mu_\gamma^{\eta'}\mathcal{L}_{\hat{g}^\gamma}\mu_m^\beta=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{L}_{\phi_{j-2}}(\nu_\beta^{\alpha_j})\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mu_m^\beta-\mathcal{O}_\gamma^{\alpha_{j-1}}\mathcal{T}_m^{\gamma,\alpha_j}=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{L}_{\phi_{j-2}}(\nu_{\beta'}^{\alpha_j})\mu_\gamma^{\beta'}\nu_\beta^\gamma\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mu_m^\beta-\mathcal{O}_\gamma^{\alpha_{j-1}}\mathcal{T}_m^{\gamma,\alpha_j}=$$

$$\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{O}_\gamma^{\alpha_j}\mathcal{T}_m^{\alpha_{j-1},\gamma}-\mathcal{O}_\gamma^{\alpha_{j-1}}\mathcal{T}_m^{\gamma,\alpha_j}$$

Hence, for this first term in (6.40) we obtain:

$$\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\eta}\mathcal{L}_{g^\beta}h_m=\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-\mathcal{O}_\gamma^{\alpha_j}\mathcal{T}_m^{\alpha_{j-1},\gamma}-\mathcal{O}_\gamma^{\alpha_{j-1}}\mathcal{T}_m^{\gamma,\alpha_j}\tag{6.41}$$

Regarding the second term in (6.40) we have:

$$-\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{g^\eta}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\beta}h_m=-\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{g^\eta}\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=-\nu_\beta^{\alpha_j}\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_\beta^{\alpha_j}\mathcal{L}_{\phi_{j-2}}\mu_m^\beta)+\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_\beta^{\alpha_j})\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}+\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_\beta^{\alpha_j})\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}+\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_{\beta'}^{\alpha_j})\mu_\gamma^{\beta'}\nu_\beta^\gamma\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}+\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_{\beta'}^{\alpha_j})\mu_\gamma^{\beta'}\nu_\beta^\gamma\mathcal{L}_{\phi_{j-2}}\mu_m^\beta-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\nu_{\beta'}^{\alpha_j}\mu_\gamma^{\beta'})\nu_\beta^\gamma\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}-\nu_{\beta'}^{\alpha_j}\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}(\mu_\gamma^{\beta'})\nu_\beta^\gamma\mathcal{L}_{\phi_{j-2}}\mu_m^\beta=$$

$$-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}-\mathcal{T}_\gamma^{\alpha_{j-1},\alpha_j}\mathcal{P}_m^\gamma$$

Hence, for this second term in (6.40) we obtain:

$$-\nu_\beta^{\alpha_j}\nu_\eta^{\alpha_{j-1}}\mathcal{L}_{g^\eta}\mathcal{L}_{\phi_{j-2}}\mathcal{L}_{g^\beta}h_m=-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}\mathcal{P}_m^{\alpha_j}-\mathcal{T}_\gamma^{\alpha_{j-1},\alpha_j}\mathcal{P}_m^\gamma\tag{6.42}$$

By substituting (6.41) and (6.42) in (6.40) and by reminding the reader that in our setting the components of th etensor $\mathcal{T}$ that have the lower index equal to zero vanish, we immediately obtain (6.39).

■

$$\mathcal{L}_{\phi_j^{\alpha_{j-1},\alpha_j}}h_m=\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-O_k^{\alpha_j}\mathcal{T}_m^{\alpha_{j-1},k}-O_k^{\alpha_{j-1}}\mathcal{T}_m^{k,\alpha_j}-\mathcal{L}_{\hat{g}^{\alpha_{j-1}}}P_m^{\alpha_j}-\mathcal{T}_k^{\alpha_{j-1},\alpha_j}P_m^k-\mathcal{L}_{\hat{g}^{\alpha_j}}\mathcal{L}_{\phi_{j-1}}h_m$$

$$\mu_m^{\alpha_{j-1},\alpha_j}=d_x\mathcal{L}_{\phi_{j-2}}\mathcal{T}_m^{\alpha_{j-1},\alpha_j}-O_k^{\alpha_j}d_x\mathcal{T}_m^{\alpha_{j-1},k}-O_k^{\alpha_{j-1}}d_x\mathcal{T}_m^{k,\alpha_j}-P_m^kd_x\mathcal{T}_k^{\alpha_{j-1},\alpha_j}$$

Lemma 18 *In general, it exists a finite $m\le n+2$ such that $\mathcal{D}\mathcal{T}_k^{\alpha,\beta}\in\Omega_m,\ \forall\alpha,\beta,k$.*

Proof: From equation (6.37) we have:

$$\mathcal{L}_{\phi_{j+2}^{\cdots,\alpha_{j+1},\alpha_{j+2}}} h_m = \tag{6.43}$$

$$\mathcal{L}_{\phi_j^{\cdots}} \mathcal{T}_m^{\alpha_{j+1},\alpha_{j+2}} - \mathcal{O}_k^{\cdots,\alpha_{j+2}} \mathcal{T}_m^{\alpha_{j+1},k} - \mathcal{O}_k^{\cdots,\alpha_{j+1}} \mathcal{T}_m^{k,\alpha_{j+2}} + \mathcal{O}_m^{\cdots,k} \mathcal{T}_k^{\alpha_{j+1},\alpha_{j+2}} - \mathcal{L}_{\hat{g}^{\alpha_{j+1}}} \mathcal{P}_m^{\cdots,\alpha_{j+2}} - \mathcal{L}_{\hat{g}^{\alpha_{j+2}}} \mathcal{L}_{\phi_{j+1}^{\cdots,\alpha_{j+1}}} h_m$$

Let us introduce the following two tensors, for a given integer j:

- ${}^j\mathcal{Z}_m^{\cdots,\alpha_{j+1},\alpha_{j+2}} \triangleq \mathcal{L}_{\phi_{j+2}^{\cdots,\alpha_{j+1},\alpha_{j+2}}} h_m$;
- ${}^j\mathcal{B}_m^{\cdots,\alpha_{j+1},\alpha_{j+2}} \triangleq \mathcal{L}_{\phi_j^{\cdots}} \mathcal{T}_m^{\alpha_{j+1},\alpha_{j+2}}$;

Both of them have $j+2$ upper indexes (Greek) and one lower index (Latin). By construction, the differentials of all the $(m_w+1)^{j+2}m_w$ components of the tensor ${}^j\mathcal{Z} \in \Omega_{j+3}$. On the other hand, from equation (6.43), we immediately obtain:

$$\begin{aligned} \mathcal{D}({}^j\mathcal{Z}_m^{\cdots,\alpha_{j+1},\alpha_{j+2}}) = \mathcal{D}({}^j\mathcal{B}_m^{\cdots,\alpha_{j+1},\alpha_{j+2}}) - \mathcal{D}\mathcal{O}_k^{\cdots,\alpha_{j+2}} \mathcal{T}_m^{\alpha_{j+1},k} - \mathcal{O}_k^{\cdots,\alpha_{j+2}} \mathcal{D}\mathcal{T}_m^{\alpha_{j+1},k} \\ -\mathcal{D}\mathcal{O}_k^{\cdots,\alpha_{j+1}} \mathcal{T}_m^{k,\alpha_{j+2}} - \mathcal{O}_k^{\cdots,\alpha_{j+1}} \mathcal{D}\mathcal{T}_m^{k,\alpha_{j+2}} + \mathcal{D}\mathcal{O}_m^{\cdots,k} \mathcal{T}_k^{\alpha_{j+1},\alpha_{j+2}} + \mathcal{O}_m^{\cdots,k} \mathcal{D}\mathcal{T}_k^{\alpha_{j+1},\alpha_{j+2}} \\ -\mathcal{D}\mathcal{L}_{\hat{g}^{\alpha_{j+1}}} \mathcal{P}_m^{\cdots,\alpha_{j+2}} - \mathcal{D}\mathcal{L}_{\hat{g}^{\alpha_{j+2}}} \mathcal{L}_{\phi_{j+1}^{\cdots,\alpha_{j+1}}} h_m \end{aligned} \tag{6.44}$$

By using lemma 16 we obtain the following results:

- $-\mathcal{D}\mathcal{O}_k^{\cdots,\alpha_{j+2}} \mathcal{T}_m^{\alpha_{j+1},k} - \mathcal{D}\mathcal{O}_k^{\cdots,\alpha_{j+1}} \mathcal{T}_m^{k,\alpha_{j+2}} + \mathcal{D}\mathcal{O}_m^{\cdots,k} \mathcal{T}_k^{\alpha_{j+1},\alpha_{j+2}} \in \Omega_{j+2}$;
- $-\mathcal{D}\mathcal{L}_{\hat{g}^{\alpha_{j+1}}} \mathcal{P}_m^{\cdots,\alpha_{j+2}} \in \Omega_{j+3}$.

Additionally, $-\mathcal{D}\mathcal{L}_{\hat{g}^{\alpha_{j+2}}} \mathcal{L}_{\phi_{j+1}^{\cdots,\alpha_{j+1}}} h_m \in \Omega_{j+3}$. Hence, from (6.44), we obtain that the following covector:

$$\begin{aligned} {}^j\mathcal{Z}_m'^{\cdots,\alpha_{j+1},\alpha_{j+2}} = \\ \mathcal{D}({}^j\mathcal{B}_m^{\cdots,\alpha_{j+1},\alpha_{j+2}}) - \mathcal{O}_k^{\cdots,\alpha_{j+2}} \mathcal{D}\mathcal{T}_m^{\alpha_{j+1},k} - \mathcal{O}_k^{\cdots,\alpha_{j+1}} \mathcal{D}\mathcal{T}_m^{k,\alpha_{j+2}} + \mathcal{O}_m^{\cdots,k} \mathcal{D}\mathcal{T}_k^{\alpha_{j+1},\alpha_{j+2}} \end{aligned} \tag{6.45}$$

belongs to Ω_{j+3}. We proceed as in the case $m_w = 1$, $g^0 = 0$. Let us denote by j^* the smallest integer such that the differentials of all the $(m_w+1)^{j^*+2}m_w$ components of the tensor ${}^{j^*}\mathcal{B}$ can be expressed as linear combinations of the differentials of the components of all the tensors ${}^j\mathcal{B}$, $j = 0, 1, \cdots, j^*-1$.

$$\mathcal{D}({}^{j^*}\mathcal{B}_m^{\cdots,\alpha_{j^*+1},\alpha_{j^*+2}}) = \sum_{j=0}^{j^*-1} {}^j c^k_{\cdots,\beta_{j+1},\beta_{j+2}} \mathcal{D}({}^j\mathcal{B}_k^{\cdots,\beta_{j+1},\beta_{j+2}}) + c^k \mathcal{D}h_k \tag{6.46}$$

where the dummy indexes k, $\beta_1, \cdots, \beta_{j+2}$ are summed up, according to the Einstein notation. Note that j^* is a finite integer and in particular $j^* \leq n-1$. Indeed, if this would not be the case, the dimension of the codistribution generated by $\mathcal{D}h$ and the differentials of all the components of the tensors ${}^j\mathcal{B}$, $j = 0, \cdots, n-1$, would be $n+1$, i.e., larger than n. From (6.46) and (6.45) we obtain:

$${}^{j^*}\mathcal{Z}_m'^{\cdots,\alpha_{j^*+1},\alpha_{j^*+2}} = \sum_{j=0}^{j^*-1} {}^j c^k_{\cdots,\beta_{j+1},\beta_{j+2}} \mathcal{D}({}^j\mathcal{B}_k^{\cdots,\beta_{j+1},\beta_{j+2}}) + c^k \mathcal{D}h_k \tag{6.47}$$

$$-\mathcal{O}_k^{\cdots,\alpha_{j^*+2}}\mathcal{D}\mathcal{T}_m^{\alpha_{j^*+1},k} - \mathcal{O}_k^{\cdots,\alpha_{j^*+1}}\mathcal{D}\mathcal{T}_m^{k,\alpha_{j^*+2}} + \mathcal{O}_m^{\cdots,k}\mathcal{D}\mathcal{T}_k^{\alpha_{j^*+1},\alpha_{j^*+2}}$$

From equation (6.45), for $j = 0, \cdots, j^* - 1$, we obtain:

$$\mathcal{D}({}^j\mathcal{B}_k^{\cdots,\beta_{j+1},\beta_{j+2}}) ={}^j \mathcal{Z}_k'^{\cdots,\beta_{j+1},\beta_{j+2}} + \mathcal{O}_l^{\cdots,\beta_{j+2}}\mathcal{D}\mathcal{T}_k^{\beta_{j+1},l} + \mathcal{O}_l^{\cdots,\beta_{j+1}}\mathcal{D}\mathcal{T}_k^{l,\beta_{j+2}} - \mathcal{O}_k^{\cdots,l}\mathcal{D}\mathcal{T}_l^{\beta_{j+1},\beta_{j+2}}$$

By substituting in (6.47) we obtain:

$$\begin{aligned}
&{}^{j^*}\mathcal{Z}_m'^{\cdots,\alpha_{j^*+1},\alpha_{j^*+2}} - c^k\mathcal{D}h_k - \sum_{j=0}^{j^*-1} {}^jc^k_{\cdots,\beta_{j+1},\beta_{j+2}}\, {}^j\mathcal{Z}_k'^{\cdots,\beta_{j+1},\beta_{j+2}} = \qquad (6.48)\\
&\sum_{j=0}^{j^*-1} {}^jc^k_{\cdots,\beta_{j+1},\beta_{j+2}}\{\mathcal{O}_l^{\cdots,\beta_{j+2}}\mathcal{D}\mathcal{T}_k^{\beta_{j+1},l} + \mathcal{O}_l^{\cdots,\beta_{j+1}}\mathcal{D}\mathcal{T}_k^{l,\beta_{j+2}} - \mathcal{O}_k^{\cdots,l}\mathcal{D}\mathcal{T}_l^{\beta_{j+1},\beta_{j+2}}\}\\
&-\mathcal{O}_k^{\cdots,\alpha_{j^*+2}}\mathcal{D}\mathcal{T}_m^{\alpha_{j^*+1},k} - \mathcal{O}_k^{\cdots,\alpha_{j^*+1}}\mathcal{D}\mathcal{T}_m^{k,\alpha_{j^*+2}} + \mathcal{O}_m^{\cdots,k}\mathcal{D}\mathcal{T}_k^{\alpha_{j^*+1},\alpha_{j^*+2}}
\end{aligned}$$

Since this equality holds $\forall m, \alpha_1, \cdots, \alpha_{j^*+2}$, (6.48) consists of $m_w(m_w+1)^{j^*+2}$ equations. We remark that the left hand side of these equations consists of the sum of covectors that belong to Ω_{j^*+3}. In the right-hand side we have the differentials of all the $m_w(m_w+1)^2$ components of the tensor $\mathcal{T}$. In general, (6.48) can be used to express these last differentials in terms of covectors that belong to Ω_{j^*+3}. Therefore, by setting $m \triangleq j^* + 3$, we have $m \leq n + 2$ and $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m, \forall \alpha, \beta, k$ ■

The previous lemma ensures that, in general, it exists a finite $m \leq n+2$ such that $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m, \forall \alpha, \beta, k$. Note that the previous proof holds if the matrix that expresses the dependency of (6.48) on the terms $\mathcal{D}\mathcal{T}_k^{\alpha,\beta}$ can be inverted. This holds in general, with the exception of the trivial case considered in lemma 14.

The following theorem allows us to obtain the criterion to stop algorithm 5:

Theorem 4 *If $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m$, $\forall \alpha, \beta, k$, and $\Omega_{m+1} = \Omega_m$ (namely, Ω_m is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\hat{g}^\alpha}$, $\forall \alpha$, simultaneously) then $\Omega_{m+p} = \Omega_m$ $\forall p \geq 0$*

Proof: We proceed by induction. Obviously, the equality holds for $p = 0$.
Inductive step: let us assume that $\Omega_{m+p} = \Omega_m$ and let us prove that $\Omega_{m+p+1} = \Omega_m$. We have to prove that $\mathcal{D}\mathcal{L}_{\phi_{m+p}^{\alpha_1,\cdots,\alpha_{m+p}}} h \in \Omega_m$, $\forall \alpha_1, \cdots, \alpha_{m+p}$. Indeed, from the inductive assumption, we know that $\Omega_{m+p} (= \Omega_m)$ is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\hat{g}^\alpha}$, $\forall \alpha$. Additionally, because of this invariance, by using proposition 13, we obtain that Ω_m is also invariant under $\mathcal{L}_{\phi_j^{\alpha_1,\cdots,\alpha_j}}$, $\forall \alpha_1, \cdots, \alpha_j$, for $j = 1, 2, \cdots, m+p-1$. Since $\mathcal{D}\mathcal{T}_k^{\alpha,\beta} \in \Omega_m$, $\forall \alpha, \beta, k$, by computing the differential of equation equation (6.37) for $j = m+p$, it is immediate to obtain that $\mathcal{D}\mathcal{L}_{\phi_{m+p}^{\alpha_1,\cdots,\alpha_{m+p}}} h \in \Omega_m$, $\forall \alpha_1, \cdots, \alpha_{m+p}$ ■

We conclude this section by providing an upper bound for the number of steps that are in general necessary to achieve the convergence. The dimension of Ω_{j^*+2} is at least the dimension of the span of the covectors: $\mathcal{D}h$, ${}^0\mathcal{Z}_k'^{\alpha_1,\alpha_2}$, ${}^1\mathcal{Z}_k'^{\alpha_1,\alpha_2,\alpha_3}$, $\cdots$, ${}^{j^*-1}\mathcal{Z}_k'^{\alpha_1,\cdots,\alpha_{j^*+1}}$. From the definition of j^*, we know that among the vectors $\mathcal{D}h$, $\mathcal{D}\,{}^0\mathcal{B}_k^{\alpha_1,\alpha_2}$, $\mathcal{D}\,{}^1\mathcal{B}_k^{\alpha_1,\alpha_2,\alpha_3}$, $\cdots$, $\mathcal{D}\,{}^{j^*-1}\mathcal{B}_k^{\alpha_1,\cdots,\alpha_{j^*+1}}$ at least j^*+1 are independent meaning that the dimension of their span is at least j^*+1. Hence, from (6.45), it easily follows that the dimension of the span of the vectors $\mathcal{D}h$, ${}^0\mathcal{Z}_k'^{\alpha_1,\alpha_2}$, ${}^1\mathcal{Z}_k'^{\alpha_1,\alpha_2,\alpha_3}$, $\cdots$, ${}^{j^*-1}\mathcal{Z}_k'^{\alpha_1,\cdots,\alpha_{j^*+1}}$, $\mathcal{D}\mathcal{T}_k^{\alpha,\beta}$ is at least j^*+1. Since Ω_{j^*+3} contains this span, its dimension is at least j^*+1. Therefore, the condition $\Omega_{m+1} = \Omega_m$, for $m \geq j^*+3$ is achieved for $m \leq n+2$.

6.3.3 Extension to the case of multiple known inputs

It is immediate to repeat all the steps carried out in the previous two subsections and extend the validity of theorem 3 to the case of multiple known inputs ($m_u > 1$). Additionally, also theorem 4 can be easily extended to cope with the case of multiple known inputs. In this case, requiring that $\Omega_{m+1} = \Omega_m$ means that Ω_m must be invariant with respect to all $\mathcal{L}_{\hat{g}^\alpha}$ and all $\mathcal{L}_{f^i}$ simultaneously.

Chapter 7

Conclusion

The goal of this book was to provide the analytic solution of a fundamental open problem in control theory. This problem, called the Unknown Input Observability (UIO) problem, was introduced in the seventies. It consists in providing the analytic condition to obtain the state observability of a nonlinear systems driven by both known and unknown inputs. This book provided and illustrated this analytic solution that can be considered the natural extension of the observability rank condition in [18] to account for the presence of unknown inputs.

The condition is based on the computation of a codistribution (the observable codistribution). As in the standard case of only known inputs, the observable codistribution is obtained by a recursive algorithm (algorithm 5). With respect to the algorithm that generates the observable codistribution in the case of only known inputs, in the new algorithm, the vector fields that correspond to the unknown inputs $w_1, \cdots, w_{m_w}$ (i.e., $g^1(x), \cdots, g^{m_w}(x)$) and the drift in the dynamics (i.e., $g^0(x)$) must be substituted with the vector fields $\hat{g}^0(x), \hat{g}^1(x), \cdots, \hat{g}^{m_w}(x)$ defined in (3.9). Additionally, the recursive step of the new algorithm also contains the differentials of the Lie derivatives of the scalar functions $h_1(x), \cdots, h_{m_w}(x)$ defined in appendix A, along a new set of vector fields (${}^i\phi_j^{\alpha_1,\cdots,\alpha_j}$). These vector fields are obtained starting from the vector fields that characterize the system dynamics and by performing with them the operation defined by definition 1, which is a new abstract operation that extends the Lie bracket. Specifically, it is the Lie bracket along a set of vector fields. In practice, the entire observable codistribution was obtained by a very simple recursive algorithm (algorithm 5).

Finally, we proved that this recursive algorithm converges in a finite number of steps and we provided the analytic criterion to establish that the convergence has been reached. In particular, the convergence is attained in at most $n+2$ steps, where n is the state dimension.

Very surprisingly, the complexity of this algorithm is comparable to the complexity of the standard algorithm to compute the observable codistribution in the case without unknown inputs (i.e., the observability rank condition). Given any nonlinear system characterized by any type of nonlinearity, driven by both known and unknown inputs, the state observability is obtained automatically, i.e., without human intervention (e.g., by the usage of a very simple code that uses symbolic computation). This is a fundamental practical (and unexpected) advantage.

On the other hand, the analytic derivations and all the proofs necessary to analytically derive the algorithm and its convergence properties and to prove their general validity are very complex and they are extensively based on an ingenious analogy with the theory of General Relativity. In practice, these derivations largely use Ricci calculus with tensors (in particular, we largely adopted the Einstein notation to achieve notational brevity).

The analytic criterion here derived is the follow up of the analytic criterion given in [43], that

is the analytic solution of the UIO problem for nonlinear driftless systems in presence of a single unknown input. The new general analytic criterion accounts for the presence of a drift and the presence of multiple unknown inputs. Very interestingly, in the aforementioned analogy with the theory of General Relativity, the presence of the drift term corresponds to the presence of a time dimension in relativity and the presence of unknown inputs corresponds to the space dimension in relativity (i.e., in the aforementioned analogy, the space dimension of relativity equals the number of unknown inputs). In this sense, the solution in [43], which holds for driftless systems and with a single unknown input, corresponds to the trivial case of a space-time frozen with respect to time and with a single spatial dimension (this is the reason why the derivation of the solution in [43] did not require the use of Ricci calculus).

The analytic criterion was illustrated by checking the observability of several nonlinear systems driven by multiple known inputs and multiple unknown inputs, ranging from planar robotics up to advanced nonlinear systems. In particular, the last applications were in the framework of visual-inertial sensor fusion (sections 5.6 and 5.7). For this problem, the application of the analytic criterion provided a remarkable and amazing result, which could have relevance on the problem of visual-vestibular integration for self-motion perception in neuroscience.

We conclude by emphasizing that, compared with existing methods, the analytic condition proposed in this book, has two fundamental novelties:

- It holds for any type of nonlinearity.
- It can be implemented automatically.

Regarding the first novelty, note that the analytic condition holds for any set of vector fields $f^i(x)$ $(i = 1, \cdots, m_u)$, $g^j(x)$ $(j = 1, \cdots, m_w)$ and $g^0(x)$ in (2.1). Additionally, it holds for any set of outputs $h_i(x)$ $(i = 1, \cdots, p)$ in (A.1). In other words, the solution can be applied independently of the complexity and type of nonlinearity of the aforementioned functions. The solutions available in the literature refer always to very specific cases of functions $f^i(x)$ and/or $g^0(x)$ and/or $g^j(x)$ and/or $h_i(x)$.

Regarding the second novelty, note that, the procedure summarized in chapter 4, can be implemented automatically, e.g., by using a symbolic toolbox, without human intervention. It suffices to simply follow the steps in the procedure of chapter 4, independently of the system complexity, state dimension, type of nonlinearity. To the best of our knowledge, the existing conditions in literature to check the state observability cannot be implemented automatically, with the exception of the observability rank condition (but this does not account for the presence of unknown inputs).

Appendix A

Canonic form with respect to the unknown inputs

We refer to the system characterized by the following equations:

$$\begin{cases} \dot{x} = g^0(x) + \sum_{i=1}^{m_u} f^i(x) u_i + \sum_{j=1}^{m_w} g^j(x) w_j \\ y_k = h_k(x), \quad k = 1, \cdots, p \end{cases} \tag{A.1}$$

We introduce the following notation:

- $\mathcal{F}$ is the space of functions that contains all the outputs $h_1, \cdots, h_p$ and their Lie derivatives along the vector fields $f^1, \cdots, f^{m_u}$, up to any order;
- $\mathcal{DF}$ is the codistribution generated by the differentials, with respect to the state x, of the functions in $\mathcal{F}$;
- $\mathcal{L}_G\mathcal{F}$ is the space of functions that consists of the Lie derivatives along G of the functions in $\mathcal{F}$ (we remind the reader that G is the vector defined in (6.3));
- $\mathcal{D}_w\mathcal{L}_G\mathcal{F}$ is the codistribution generated by the differentials, with respect to the unknown input vector w, of the functions in $\mathcal{L}_G\mathcal{F}$.

It is immediate to prove the following properties:

- From $\mathcal{F}$ we can select a set of functions such that their differentials generate $\mathcal{DF}$.
- It exists an integer m such that, by running m recursive steps of algorithm 6, we obtain a codistribution $\bar{\Omega}_m$ that contains $[\mathcal{DF},\ 0_{mm_w}]$ (i.e., the codistribution $\mathcal{DF}$ once embedded in $\mathbb{R}^{n+mm_w}$)[1].
- In general, the functions belonging to $\mathcal{L}_G\mathcal{F}$, are functions of x and $w_1, \cdots, w_{m_w}$.
- The dimension of the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}$ cannot exceed the dimension of w (i.e., it cannot exceed m_w).

[1]Note that, from proposition 5, the functions in $\mathcal{F}$ are weakly observable.

We introduce the following definition:

Definition 7 (Canonic Form) *The system in (A.1) is in canonic form with respect to the unknown inputs if the dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}$ is m_w.*

We remark that, if the system is in canonic form with respect to the unknown inputs, we can select m_w functions from $\mathcal{L}_G\mathcal{F}$ whose differentials with respect to w are independent. Therefore, we can select m_w functions from $\mathcal{F}$ such that, by using these functions for the selection of $h_1, \cdots, h_{m_w}$ needed to run algorithm 5, the tensor μ, defined in (3.3), is non-singular. In the case $m_w = 1$, this means that, we can select a function from $\mathcal{F}$, such that, the quantity L^1_g defined in (3.1), does not vanish. Note that we are allowed to use the functions in $\mathcal{F}$ as system outputs, because these functions are weakly observable, as we mentioned above.

In the rest of this chapter, we show that there exists a finite number of transformations, such that, any system characterized by (A.1), can be either reduced in canonic form, or part of its unknown inputs is spurious (i.e., they do not affect the observability properties of the state). We will call the procedure that consists of these transformations, *system canonization.* For the clarity sake, we distinguish the case of $m_w = 1$ from the general case. In particular, in section A.1 we discuss the case $m_w = 1$ (both in the case $g^0 = 0$ (section A.1.1) and $g^0 \neq 0$ (section A.1.2)) and in section A.2 the general case.

A.1 System canonization in the case of a single unknown input

Let us suppose that the system is not in canonic form. From definition 7, we know that the dimension of the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}$ is 0. This means that the functions in $\mathcal{L}_G\mathcal{F}$ are independent of w. We discuss separately the cases $g^0 = 0$ and $g^0 \neq 0$.

A.1.1 $g^0 = 0$

In this case, $\mathcal{L}_G\mathcal{F}$ only contains the zero function (the Lie derivative along G of any function in $\mathcal{F}$ vanishes). As a result, any order Lie derivative of any outputs, computed along $F^1, \cdots, F^{m_u}$ (which are the vector fields defined in (6.4)) and at least once along G, vanishes. This means that, in algorithm 6, we can ignore the contribution due to the Lie derivative along G. But this trivially means that the observable codistribution is precisely $\mathcal{D}\mathcal{F}$ and the unknown input is spurious.

A.1.2 $g^0 \neq 0$

We perform a recursive procedure that consists of a finite number of steps. We set $\mathcal{F}^0 \triangleq \mathcal{F}$, $\mathcal{D}\mathcal{F}^0 \triangleq \mathcal{D}\mathcal{F}$ and $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^0 \triangleq \mathcal{D}_w\mathcal{L}_G\mathcal{F}$. The $(l+1)^{th}$ step of the procedure ($l \geq 0$) consists of the following operations:

1. Define the space of functions $\mathcal{F}^{l+1}$, as the space that contains all the functions in $\mathcal{F}^l + \mathcal{L}_G\mathcal{F}^l$ and their Lie derivatives along the vector fields $f^1, \cdots, f^{m_u}$, up to any order.
2. Build the codistribution $\mathcal{D}\mathcal{F}^{l+1}$ defined as the codistribution that contains all the differentials, with respect to the state x, of the functions in $\mathcal{F}^{l+1}$ (note that $\mathcal{D}\mathcal{F}^{l+1}$ is included in the observable codistribution of the system defined by (A.1), as long as it exists an integer m, such that, by running m recursive steps of algorithm 6, we obtain a codistribution $\bar{\Omega}_m$ that contains $[\mathcal{D}\mathcal{F}^{l+1},\ 0_{mm_w}]$).

3. Build the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}$ as the codistribution generated by the differentials with respect to w of the functions that belong to $\mathcal{L}_G\mathcal{F}^{l+1}$.

We remark that, at each step, we can have two distinct results:

- The dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}$ is 0 ($dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}) = 0$);
- The dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}$ is 1 ($dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}) = 1$).

In the latter case we stop the procedure. We select a scalar function in $\mathcal{F}^{l+1}$ such that its Lie derivative along g does not vanish (it exists because $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}) = 1$). The system is in canonic form by using this scalar function as a system output (we are allowed to use this function as an output since its differential belongs to $\mathcal{D}\mathcal{F}^{l+1}$, which is included in the observable codistribution of the system defined by (A.1)).

Let us consider the former case (i.e., $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}) = 0$). We easily have: $\mathcal{L}_g\mathcal{F}^{l+1} = \{0\}$. Hence, $\mathcal{F}^{l+2} = \mathcal{F}^{l+1}+\mathcal{L}_G\mathcal{F}^{l+1} = \mathcal{F}^{l+1}+\mathcal{L}_{g^0}\mathcal{F}^{l+1}$ We proceed as follows. We check if $\mathcal{L}_{g^0}\mathcal{D}\mathcal{F}^{l+1} \subseteq \mathcal{D}\mathcal{F}^{l+1}$. Note that this condition will be satisfied in at least $n-1$ steps. Once this condition is satisfied, it means that $\mathcal{L}_g\mathcal{L}_{g^0}\mathcal{F}^{l+1} = \{0\}$. Hence, $\mathcal{L}_g\mathcal{L}_G\mathcal{F}^{l+1} = \{0\}$ and $\mathcal{L}_g\mathcal{F}^{l+2} = \{0\}$. Therefore: $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+2}) = 0$. By induction, this means that $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+p}) = 0$ for any integer p. In this case, the unknown input is spurious and the observable codistribution is $\mathcal{D}\mathcal{F}^{l+1}$.

Algorithm 8 provides the pseudo code for the system canonization.

Algorithm 8 (Canonization $m_w = 1,\ g^0 \neq 0$**)**
if $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}) == 1$ ***then***
Set h one of the functions in $\mathcal{L}_G\mathcal{F}$ s.t. $\mathcal{L}_g h \neq 0$. The system is in canonic form. Algorithm 3 can be implemented by using this function.
RETURN
end if
while $\mathcal{L}_{g^0}\mathcal{D}\mathcal{F} \nsubseteq \mathcal{D}\mathcal{F}$ ***do***
Set $\mathcal{F} := \mathcal{F} + \mathcal{L}_G\mathcal{F}$ and close with respect to $\mathcal{L}_{f^i}$ $\forall i$
if $dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F}) == 1$ ***then***
Set h one of the functions in $\mathcal{L}_G\mathcal{F}$ s.t. $\mathcal{L}_g h \neq 0$. The system is in canonic form. Algorithm 3 can be implemented by using this function.
RETURN
end if
end while
The unknown input is spurious and the Observable codistribution is $\mathcal{D}\mathcal{F}$

A.2 System canonization in the general case

Let us suppose that the system is not in canonic form. From definition 7, we know that the dimension of the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}$ is smaller than m_w. As in the case $m_w = 1,\ g^0 \neq 0$, we introduce a recursive procedure.

A.2.1 The recursive procedure to perform the system canonization in the general case

We set $\mathcal{F}^0 \triangleq \mathcal{F}$, $\mathcal{DF}^0 \triangleq \mathcal{DF}$, ${}^0w \triangleq w$ and $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^0 \triangleq\ \mathcal{D}_w\mathcal{L}_G\mathcal{F}$. In addition, we denote by d^0 the dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^0$ ($d^0 < m_w$). The $(l+1)^{th}$ step of this procedure ($l \geq 0$) consists of 6 operations. In particular, the last three operations are the three operations that characterize the $(l+1)^{th}$ step of the recursive procedure introduced in the case $m_w = 1,\ g^0 \neq 0$. We provide first the list of the operations and after we detail them. Note that, at each step, the state will be in general augmented and the vector fields, $f^1, \cdots, f^{m_u}$, will be modified, accordingly (i.e., the modified vectors will characterize the dynamics of the augmented state). The six operations are:

1. Redefine the unknown inputs in such a way that all the functions in $\mathcal{L}_G\mathcal{F}^l$ depend only on x and d^l unknown inputs (i.e., they are independent of the remaining $m_w - d^l$ unknown inputs). In section A.2.1 we detail the method needed to obtain this result. The new unknown input vector will be denoted by $\tilde{w}$. The functions in $\mathcal{L}_G\mathcal{F}^l$ depend only on x and the first d^l entries of $\tilde{w}$.

2. Augment the state by including in it the first d^l entries of $\tilde{w}$. Now the functions in $\mathcal{L}_G\mathcal{F}^l$ depend only on the augmented state.

3. Define the new vector of unknown inputs (${}^{l+1}w$) as follows. Its first d^l entries are the first order time derivatives of the entries of $\tilde{w}$. The last $m_w - d^l$ coincide with the last $m_w - d^l$ entries of $\tilde{w}$. In other words: ${}^{l+1}w \triangleq [\dot{\tilde{w}}_1, \cdots, \dot{\tilde{w}}_{d^l},\ \tilde{w}_{d^l+1}, \cdots, \tilde{w}_{m_w}]$.

4. Define the space of functions $\mathcal{F}^{l+1}$, as the space that contains all the functions in $\mathcal{F}^l + \mathcal{L}_G\mathcal{F}^l$ and their Lie derivatives along the vector fields $f^1, \cdots, f^{m_u}$, up to any order. Note that the vector fields $f^1, \cdots, f^{m_u}$ are now the ones that characterize the dynamics of the augmented state.

5. Build the codistribution $\mathcal{DF}^{l+1}$ defined as the smallest codistribution that contains the differentials with respect to the augmented state of the functions in $\mathcal{F}^l + \mathcal{L}_G\mathcal{F}^l$ and it is invariant with respect to the Lie derivatives along the vector fields $f^1, \cdots, f^{m_u}$. In other words, $\mathcal{DF}^{l+1}$ is the codistribution generated by the differentials with respect to the augmented state of the functions in $\mathcal{F}^{l+1}$ (note that $\mathcal{DF}^{l+1}$ is included in the observable codistribution of the system defined by (A.1)).

6. Build the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}$ as the codistribution generated by the differentials with respect to ${}^{l+1}w$ of the functions that belong to $\mathcal{L}_G\mathcal{F}^{l+1}$.

First operation

Since the dimension of $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^l$ is d^l, we can select d^l functions in $\mathcal{L}_G\mathcal{F}^l$ whose differentials with respect to ${}^lw$ are independent. Let us denote them by ${}^l\gamma_1, \cdots, {}^l\gamma_{d^l}$. We also denote by ${}^lh_i$ ($i =$

$1, \cdots, d^l$) the function in $\mathcal{F}^l$ such that ${}^l\gamma_i = \mathcal{L}_G\ {}^l h_i$. We have: ${}^l\gamma_i = \mathcal{L}_G\ {}^l h_i = \mathcal{L}_{g^0}\ {}^l h_i + \mathcal{L}_{g^j}\ {}^l h_i\ w_j$. By denoting ${}^l\gamma_i^j \triangleq \mathcal{L}_{g^j}\ {}^l h_i$ we have:

$$ {}^l\gamma_i = \mathcal{L}_{g^0}\ {}^l h_i + {}^l\gamma_i^j w_j \tag{A.2} $$

Since the differentials with respect to ${}^l w$ of ${}^l\gamma_1, \cdots, {}^l\gamma_{d^l}$ are independent, we can extract from ${}^l\gamma_i^j$ a non singular two index tensor, whose indexes take the values $1, \cdots, d^l$. Additionally, let us reorder the unknown inputs in such a way that this tensor coincide with ${}^l\gamma_i^j$ for $j = 1, \cdots, d^l$. We denote this tensor by ${}^l\mu_i^j$. We can write (A.2) as follows:

$$ {}^l\gamma_i = \mathcal{L}_{g^0}\ {}^l h_i + \sum_{j=1}^{d^l} {}^l\mu_i^j w_j + \sum_{j=d^l+1}^{m_w} {}^l\gamma_i^j w_j \tag{A.3} $$

We denote by ${}^l\nu$ the inverse of ${}^l\mu$. We introduce the following coordinate change (in the unknown inputs):

$$ \begin{aligned} & w_j \to \tilde{w}_j \triangleq w_j + \sum_{i=1}^{d^l} \sum_{k=d^l+1}^{m_w} {}^l\nu_j^i\ {}^l\gamma_i^k w_k && j = 1, \cdots, d^l \\ & w_j \to \tilde{w}_j \triangleq w_j && j = d^l + 1, \cdots, m_w \end{aligned} \tag{A.4} $$

Note that this coordinate change corresponds to a redefinition of the vector fields g^j. Specifically we have:

$$ \begin{aligned} & g^j \to g^j && j = 1, \cdots, d^l \\ & g^j \to g^j - \sum_{i=1}^{d^l} \sum_{k=1}^{d^l} {}^l\nu_k^i\ {}^l\gamma_i^j g^k && j = d^l + 1, \cdots, m_w \end{aligned} \tag{A.5} $$

By an explicit computation it is possible to verify that, after this change, the functions in $\mathcal{L}_G\mathcal{F}^l$ only depend on $\tilde{w}_1, \cdots, \tilde{w}_{d^l}$, namely, on the first d^l entries of the new unknown input vector $\tilde{w}$.

Second operation

We include the first d^l entries of $\tilde{w}$ in the state, i.e.,

$$ x \to [x^T, \tilde{w}_1, \cdots, \tilde{w}_{d^l}]^T \tag{A.6} $$

Third operation

We define the new unknown input vector:

$$ {}^{l+1}w \triangleq [\dot{\tilde{w}}_1, \cdots, \dot{\tilde{w}}_{d^l},\ \tilde{w}_{d^l+1}, \cdots, \tilde{w}_{m_w}] \tag{A.7} $$

We denote by $\mathcal{D}_{{}^{l+1}w}$ the differential with respect to ${}^{l+1}w$.

Fourth operation

We define the space of functions $\mathcal{F}^{l+1}$ as follows. This space contains $\mathcal{F}^l + \mathcal{L}_G\mathcal{F}^l$ and their Lie derivative along the vector fields $f^1, \cdots, f^{m_u}$, up to any order.

Fifth operation

We denote by $\mathcal{DF}^{l+1}$ the codistribution generated by the differentials with respect to the new state in (A.6) of the functions in $\mathcal{F}^{l+1}$. $\mathcal{DF}^{l+1}$ is the smallest codistribution that contains the differentials of $\mathcal{F}^l + \mathcal{L}_G\mathcal{F}^l$ and it is invariant with respect to the Lie derivative along the vector fields $f^1, \cdots, f^{m_u}$. Note that $\mathcal{DF}^{l+1}$ is included in the observable codistribution of the system defined by (A.1).

Sixth operation

We build the codistribution $\mathcal{D}_w\mathcal{L}_G\mathcal{F}^{l+1}$ as the codistribution generated by the differentials with respect to ${}^{l+1}w$ of the functions that belong to $\mathcal{L}_G\mathcal{F}^{l+1}$. The dimension of this codistribution is larger than or equal to d^l. On the other hand, it cannot exceed the dimension of ${}^{l+1}w$, i.e, m_w. We denote this dimension with d^{l+1} and we have $d^l \leq d^{l+1} \leq m_w$.

A.2.2 Convergence of the recursive procedure

We start by remarking that, in the case $d^{l+1} = m_w$, the canonization has been completed. Indeed, we can select m_w functions in $\mathcal{L}_G\mathcal{F}^{l+1}$ whose differentials with respect to ${}^{l+1}w$ are independent. Let us denote these functions by $\mathcal{L}_G\tilde{h}_1, \cdots, \mathcal{L}_G\tilde{h}_{m_w}$. The system is in canonic form by using $\tilde{h}_1, \cdots, \tilde{h}_{m_w}$ as system outputs (we are allowed to use these functions as outputs since their differentials belong to $\mathcal{DF}^{l+1}$, which is included in the observable codistribution of the system defined by (A.1)).

In general, at each step, we can have two distinct results:

- $d^{l+1} > d^l$.
- $d^{l+1} = d^l$.

We remark that the first case cannot occur indefinitely. Hence, each time that the second case occurs, we apply algorithm 5 to the current system by only considering the first d^{l+1} unknown inputs. Let us denote by Ω^* the codistribution provided by this algorithm, once converged. We then continue to proceed with the procedure of above. At each step, denoted by $l + s$, $s \geq 2$, we check, first of all, if $d^{l+s} = d^{l+1}$. If it is larger, we do not check anything else and we start again with the steps of above until, at a given step l' we obtain again that $d^{l'+1} = d^{l'}$. If $d^{l+s} = d^{l+1}$ we check if $\Omega^* \subseteq \mathcal{DF}^{l+p}$. If this is the case we conclude that only d^{l+1} unknown inputs affect the observability properties and the remaining $m_w - d^{l+1}$ unknown inputs are spurious and can be ignored. Additionally, we conclude that Ω^* is the observable codistribution.

Note that, the final system, is characterized by a new state, according to the change described by (A.6), which is performed at each step. Additionally, this final system is in canonic form with respect to the d^{l+1} unknown inputs, which are related to the original unknown inputs by the change in (A.4) and (A.7).

Algorithm 9 provides the pseudo code for the system canonization in the general case.

Algorithm 9 (Canonization)
Set $\mathcal{F}$ the function space that includes all the outputs and close with respect to $\mathcal{L}_{f^i}$ $\forall i$
$d := dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F})$
if $d == m_w$ ***then***
The system is in canonic form. Select m_w functions in $\mathcal{F}$ such that their differentials (with respect to w) of their Lie derivatives along G, are independent.

    Use these m_w functions for the functions $h_1, \cdots, h_{m_w}$ to implement algorithm 5.
    RETURN
end if
Redefine w according to (A.4)
Augment the state according to (A.6)
Redefine w according to (A.7)
Set $\mathcal{F} := \mathcal{F} + \mathcal{L}_G\mathcal{F}$ and close with respect to $\mathcal{L}_{f^i}$ $\forall i$
$d_{old} := d$
$d := dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F})$
if $d == m_w$ ***then***
    The system is in canonic form. Select m_w functions in $\mathcal{F}$ such that their differentials (with respect to w) of their Lie derivatives along G, are independent.
    Use these m_w functions for the functions $h_1, \cdots, h_{m_w}$ to implement algorithm 5.
    RETURN
end if
while *1* ***do***
    if $d == d_{old}$ ***then***
        Select d functions in $\mathcal{F}$ such that their differentials (with respect to w) of their Lie derivatives along G, are independent.
        Implement algorithm 5 by ignoring the remaining $m_w - d$ unknown inputs. Specifically, use the selected d functions for the functions $h_1, \cdots, h_d$ to implement algorithm 5 with d unknown inputs. Denote by Ω^* the codistribution provided by the algorithm, once converged.
    end if
    while $d == d_{old}$ ***do***
        if $\Omega^* \subseteq \mathcal{D}\mathcal{F}$ ***then***
            The system is in canonic form only with respect to the first d unknown inputs. The remaining $m_w - d$ inputs are spurious. The observable codistribution is $\mathcal{D}\mathcal{F}$.
            RETURN
        end if
        Redefine w according to (A.4)
        Augment the state according to (A.6)
        Redefine w according to (A.7)
        Set $\mathcal{F} := \mathcal{F} + \mathcal{L}_G\mathcal{F}$ and close with respect to $\mathcal{L}_{f^i}$ $\forall i$
        $d := dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F})$
        if $d == m_w$ ***then***
            The system is in canonic form. Select m_w functions in $\mathcal{F}$ such that their differentials (with respect to w) of their Lie derivatives along G, are independent.
            Use these m_w functions for the functions $h_1, \cdots, h_{m_w}$ to implement algorithm 5.
            RETURN
        end if
    end while
    Redefine w according to (A.4)
    Augment the state according to (A.6)
    Redefine w according to (A.7)
    Set $\mathcal{F} := \mathcal{F} + \mathcal{L}_G\mathcal{F}$ and close with respect to $\mathcal{L}_{f^i}$ $\forall i$
    $d_{old} := d$
    $d := dim(\mathcal{D}_w\mathcal{L}_G\mathcal{F})$
    if $d == m_w$ ***then***

The system is in canonic form. Select m_w functions in $\mathcal{F}$ such that their differentials (with respect to w) of their Lie derivatives along G, are independent.
Use these m_w functions for the functions $h_1, \cdots, h_{m_w}$ to implement algorithm 5.
RETURN
end if
end while

A.2.3 Remarks to reduce the computation

We conclude this section by adding two important remarks that can significantly reduce the computational burden to perform the observability analysis of a system once it has been set in canonic form, by using the procedure of above.

The first remark is the following. For each step (m) of the above procedure, we remark that, the differentials of all the scalar functions that belong to $\mathcal{F}^m$, belong to the observable codistribution. Hence, in order to perform an observability analysis, we can consider any functions in $\mathcal{F}^m$ as a system output. Therefore, we have the following result:

Remark 3 *Let us suppose that the system is set in canonic form after m steps of the above procedure. Let us suppose that we can find m_w scalar functions, $h_1, \cdots, h_{m_w}$, in $\mathcal{F}^m$ that only depend on the state after $k < m$ steps. Additionally, let us suppose that the system obtained after k steps is in canonic form by using these functions as outputs. Instead of performing the observability analysis by implementing algorithm 5 on the system after m steps, we can implement algorithm 5 on the system after k steps, by using the outputs $h_1, \cdots, h_{m_w}$.*

Similarly, once we know that our system can be set in canonic form, we can run algorithm 6 and try to find functions of the original state, whose differential belongs to the codistribution generated by this algorithm. We remark that, the differentials of all these functions, belong to the observable codistribution. Hence, in order to perform an observability analysis, we can consider these functions as system outputs. Therefore, we have the following result:

Remark 4 *Let us consider m_w scalar functions that only depend on the original state. Let us suppose that their differentials belong to the codistribution generated by running algorithm 6 for a given number of steps. Additionally, let us suppose that, by using these functions as outputs, the original system is in canonic form. We are allowed to perform the observability analysis by applying algorithm 5 to the original system and by selecting these functions to run algorithm 5.*

Bibliography

[1] J.-P. Barbot, M. Fliess, T. Floquet, "An algebraic framework for the design of nonlinear observers with unknown inputs", IEEE Conf. on Decision and Control, New-Orleans, 2007

[2] J.-P. Barbot, D. Boutat, T. Floquet ,"An observation algorithm for nonlinear systems with unknown inputs", Automatica, Vol. 45, Issue 8, August 2009, Pages 1970-1974

[3] G. Basile and G. Marro. On the observability of linear, time invariant systems with unknown inputs. J. Optimization Theory Appl., 3:410–415, 1969.

[4] F. A. W. Belo, P. Salaris, and A. Bicchi, 3 Known Landmarks are Enough for Solving Planar Bearing SLAM and Fully Reconstruct Unknown Inputs, IROS 2010, Taipei, Taiwan

[5] S.P. Bhattacharyya Observer design for linear systems with unknown inputs IEEE Trans. on Aut. Control, 23 (1978)

[6] Bryson, M. and Sukkarieh, S., Observability analysis and active control for airborne SLAM, IEEE Transactions on Aerospace and Electronic Systems, Vol 44 , Issue 1, pp 261–280, January 2008

[7] A. Censi, Antonio Franchi, Luca Marchionni, and Giuseppe Oriolo. Simultaneous calibration of odometry and sensor parameters for mobile robots. IEEE Transactions on Robotics, April 2013.

[8] Chen, W. and Saif, M. (2006). Unknown input observer design for a class of nonlinear systems: an LMI approach. In IEEE American Control Conference.

[9] Darouach, M., Zasadzinski, M., Xu, S. J. (1994). Full-order observers for linear systems with unknown inputs. IEEE Transactions on Automatic Control, 39(3)

[10] C. De Persis, A. Isidori (2001). A geometric approach to nonlinear fault detection and isolation, IEEE Trans. Automatic Control, 46 (6), 853-865, 2001

[11] Floquet, T., Barbot, J. (2004). A sliding mode approach of unknown input observers for linear systems. In 43rd IEEE conference on decision and control. Atlantis, Paradise Island, Bahamas. December.

[12] Floquet, T., Edwards, C., Spurgeon, S. (2007). On sliding mode observers for systems with unknown inputs. Int. J. Adapt. Control Signal Process Vol 21, no. 89, 638–656, 2007.

[13] Guan, Y., Saif, M. (1991). A novel approach to the design of unknown input observers. IEEE Transactions on Automatic Control, 36(5).

[14] R. Guidorzi and G. Marro. On Wonham stabilizability condition in the synthesis of observers for unknown-input systems. Automatic Control, IEEE Transactions on, 16(5):499–500, oct 1971.

[15] Chao X. Guo and Stergios I. Roumeliotis, IMU-RGBD Camera 3D Pose Estimation and Extrinsic Calibration: Observability Analysis and Consistency Improvement, ICRA 2013, Karlsruhe, Germany

[16] Q.P. Ha, H. Trinh State and input simultaneous estimation for a class of nonlinear systems Automatica, 40 (2004).

[17] H. Hammouri and Z. Tmar. Unknown input observer for state affine systems: A necessary and sufficient condition. Automatica, 46(2):271–278, 2010.

[18] Hermann R. and Krener A.J., 1977, Nonlinear Controllability and Observability, Transaction On Automatic Control, AC-22(5): 728–740

[19] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, Towards consistent vision-aided inertial navigation, Int. Workshop on the Algorithmic Foundations of Robotics, Cambridge, MA, Jun, 2012.

[20] J.A. Hesch, D.G. Kottas, S.L. Bowman, and S.I. Roumeliotis, "Camera-IMU-based Localization: Observability Analysis and Consistency Improvement," International Journal of Robotics Research, 33(1), Jan. 2014, pp. 182-201.

[21] J.A. Hesch, D.G. Kottas, S.L. Bowman, and S.I. Roumeliotis, "Consistency Analysis and Improvement of Vision-aided Inertial Navigation," IEEE Transactions on Robotics, 30(1), Feb. 2014, pp. 158-176.

[22] Huang, Guoquan P., Mourikis, A.I. and Roumeliotis, S.I., Analysis and improvement of the consistency of extended Kalman filter based SLAM, IEEE International Conference on Robotics and Automation, 2008, 19-23 May 2008, Pasadena, CA.

[23] M. Hou, P.C. MÃ$\frac{1}{4}$ller Design of observers for linear systems with unknown inputs IEEE Transactions on Automatic Control, 37 (6) (1992).

[24] G. Huang, A.I. Mourikis, S.I. Roumeliotis: "Observability-based Rules for Designing Consistent EKF SLAM Estimators," International Journal of Robotics Research, 29(5), pp. 502-528, 2010.

[25] G. Huang, A.I. Mourikis, S.I. Roumeliotis: "A Quadratic-Complexity Observability-Constrained Unscented Kalman Filter for SLAM," IEEE Transactions on Robotics, 29(5), pp. 1226-1243, 2013.

[26] B. R. Hunt, T. Sauer and J. A. Yorke, Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces, BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY Volume 27, Number 2, October 1992

[27] Isidori A., Nonlinear Control Systems, 3rd ed., London, Springer Verlag, 1995.

[28] E. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach", The International Journal of Robotics Research, vol. 30, no. 4, pp. 407–430, Apr. 2011.

[29] J. Kelly and G. Sukhatme, Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration, Int. Journal of Robotics Research, vol. 30, no. 1, pp. 56–79, 2011.

[30] Koenig, D. and Mammar, S. (2001). Design of a class of reduced order unknown inputs nonlinear observer for fault diagnosis. In IEEE American Control Conference.

[31] D. Koening, B. Marx, D. Jacquet Unknown input observers for switched nonlinear discrete time descriptor system IEEE Transactions on Automatic Control, 53 (1) (2008).

[32] D. G. Kottas, J. A. Hesch, S. L. Bowman, and S. I. Roumeliotis, On the consistency of vision-aided inertial navigation, in Proc. of the Int. Symposium on Experimental Robotics, Canada, Jun 2012.

[33] Kwang Wee Lee, Wijesoma, W.S. and Javier, I.G., On the Observability and Observability Analysis of SLAM, IEEE/RSJ International Conference on Intelligent Robots and Systems, 9-15 Oct. 2006, Beijing, China.

[34] M. Li and A. I. Mourikis, Improving the accuracy of EKF-based visualinertial odometry, ICRA 2012, St. Paul, MN.

[35] M. Li, A.I. Mourikis: "Online Temporal Calibration for Camera-IMU Systems: Theory and Algorithms," International Journal of Robotics Research, 33(7), pp. 947-964, 2014.

[36] Martinelli, A., Scaramuzza, D., Siegwart, R. Automatic Self-Calibration of a Vision System during Robot Motion. IEEE International Conference on Robotics and Automation (ICRA 2006), Orlando, USA, May 2006.

[37] A. Martinelli, State Estimation Based on the Concept of Continuous Symmetry and Observability Analysis: the Case of Calibration, IEEE Transactions on Robotics, vol. 27, no. 2, pp. 239–255, 2011

[38] A. Martinelli, Vision and IMU data fusion: closed-form solutions for attitude, speed, absolute scale and bias determination, IEEE Transactions on Robotics, Volume 28 (2012), Issue 1 (February), pp 44–60.

[39] A. Martinelli, Observability Properties and Deterministic Algorithms in Visual-Inertial Structure from Motion, Found. and Trends in Robotics, Vol 3, 139–209.

[40] A. Martinelli, Minimalistic sensor design in visual-inertial structure from motion, 2015 IEEE/RSJ International Conference on Robotics and Automation, Seattle (USA).

[41] A. Martinelli, Visual-inertial structure from motion: observability vs minimum number of sensors, 2014 IEEE/RSJ International Conference on Robotics and Automation, Hong Kong.

[42] A. Martinelli, Nonlinear Unknown Input Observability: Analytical expression of the observable codistribution in the case of a single unknown input, SIAM - CT15 Symposium, 2015, Paris, France.

[43] A. Martinelli. Nonlinear Unknown Input Observability: The Analytic Solution in the case of a Single Unknown Input., arXiv:1501.02968 [math.OC]

[44] Mirzaei F.M. and Roumeliotis S.I., A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation, Transactions on Robotics, 2008, Vol. 24, No. 5, 1143–1156

[45] F.M. Mirzaei, D.G. Kottas, and S.I. Roumeliotis, 3D Lidar-Camera Intrinsic and Extrinsic Calibration: Observability Analysis and Analytical Least Squares-based Initialization, International Journal of Robotics Research, Special Issue on Robot Vision, 31(4), Apr. 2012, pp. 452-467.

[46] G. Panahandeh, C. X. Guo, M. Jansson, and S. I. Roumeliotis, "Observability Analysis of a Vision-Aided Inertial Navigation System using Planar Features on the Ground", in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4187-4194, Tokyo, Japan, Nov. 3-7, 2013.

[47] Perera, L.D.L., Melkumyan, A. and Nettleton, E., On the linear and nonlinear observability analysis of the SLAM problem, IEEE International Conference on Mechatronics, 14-17 April 2009, Malaga.

[48] N. Trawny, G. Huang, A.I. Mourikis, S.I. Roumeliotis: "Observability-based Consistent EKF Estimators for Multi-robot Cooperative Localization," Autonomous Robots, 30(1), pp. 99-122, 2011.

[49] Vidal-Calleja, T., Bryson, M., Sukkarieh, S., Sanfeliu, Alberto, Andrade-Cetto, J., On the Observability of Bearing-only SLAM, EEE International Conference on Robotics and Automation, 10-14 April 2007, Roma.

[50] S.H. Wang, E.J. Davison, P. Dorato Observing the states of systems with unmeasurable disturbance IEEE Transactions on Automatic Control, 20 (1975)

[51] Wang Z. and Dissanayake, G., Observability analysis of SLAM using fisher information matrix, International Conference on Control, Automation, Robotics and Vision, 2008, Hanoi.

[52] F. Yang, R.W. Wilde Observer for linear systems with unknown inputs IEEE Transactions on Automatic Control, 33 (7) (1988).